

```

RN 1 MSKGEELFTGVVPILVELDGVNKHKFSVSGEGDADYGLKIKFTCTTGKLPVWPPTL 60
RP STRAIN=GFPxm191uv;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RT "Colorful mutants of green fluorescent protein from Aequorea
RT macrodactyla.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435430; AAL33915.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 27002 MW; BD5BA2982264C018 CRC64;

Query Match 83.7%; Score 1059; DB 5; Length 238;
Best Local Similarity 80.7%; Pred. No. 3.4e-83;
Matches 192; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPILVELDGVNKHKFSVSGEGDADYGLKIKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPILVELDGVNKHKFSVSGEGDADYGLKIKFTCTTGKLPVWPPTL 60

QY 61 VTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLV 120
Db 61 VTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLV 120

QY 61 VTTLSYGILCFARYPEMKNDFFKSAPEGYVQERTIFFQDDGKYKTRGEVKEGDTLV 120
Db 61 VTTLSYGILCFARYPEMKNDFFKSAPEGYVQERTIFFQDDGKYKTRGEVKEGDTLV 120

QY 121 NRLEKGDIFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRLEKGMDFKEDGNILGHKLEYNFISHNVYIMPDKANGLKVNFKIRHNIEGGVQLAD 180

QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSQPKPKRDHVMVLEFVTAAGITHGMDELYK 238
Db 181 HYQNTVPLGDPVLLIPINHYLSQTAISKORNETRDHVMVLEFFSACGHTGMDLYK 238

RESULT 14
Q8WTC5 PRELIMINARY; PRT; 238 AA.
ID Q8WTC5
AC Q8WTC5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orange fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Chnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OC NCBI_TaxID=147615;
RN 1
RP STRAIN=OPFxm;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RT "Colorful mutants of green fluorescent protein from Aequorea
RT macrodactyla.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435432; AAL33917.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 27018 MW; 75521EA5534E573A CRC64;

Query Match 83.6%; Score 1057; DB 5; Length 238;
Best Local Similarity 81.5%; Pred. No. 5e-83;
Matches 194; Conservative 15; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPILVELDGVNKHKFSVSGEGDADYGLKIKFTCTTGKLPVWPPTL 60

```

```

Db 1 MSKGEELFTGVVPILVELDGVNKHKFSVSGEGDADYGLKIKFTCTTGKLPVWPPTL 60
QY 61 VTFSGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLV 120
Db 61 VTTLYGILCFARYPEMKNDFFKSAPEGYVQERTIFFQDDGKYKTRGEVKEGDTLV 120
QY 121 NRLEKGDIFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRLEKGMDFKEDGNILGHKLEYNFISHNVYIMPDKANGLKVNFKIRHNIEGGVQLAD 180
QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSQPKPKRDHVMVLEFVTAAGITHGMDELYK 238
Db 181 HYQNTVPLGDPVLLIPINHYLSQTAISKORNETRDHVMVLEFFSACGHTGMDLYK 238

RESULT 15
Q95UA7 PRELIMINARY; PRT; 225 AA.
ID Q95UA7
AC Q95UA7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cyan fluorescent protein (Fragment).
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastraea.
OC NCBI_TaxID=63558;
RN 1
RP SEQUENCE FROM N.A.
RA Falkowski P.G., Sun Y.;
RT "Montastraea cavernosa fluorescent protein.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY056460; AAL17905.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
DR PRODOM; PD013756; Green_fl_protein; 1.
FT NON TER 225
SQ SEQUENCE 225 AA; 25775 MW; 52DE2F716D083524 CRC64;

Query Match 19.6%; Score 247.5; DB 5; Length 225;
Best Local Similarity 25.3%; Pred. No. 2.4e-13;
Matches 65; Conservative 46; Mismatches 71; Indels 69; Gaps 9;

QY 11 VVPILVELDGVNKHKFSVSGEGDADYGLKIKFTCTTGKLPVWPPTLVTFSGVQ 69
Db 7 VMKILRMDDGIVNGHKFMITGEGKPFEGTHIILKVEGGPLPAYDILITTAFOYGNR 66
QY 70 CFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLVNRIELKGD 129
Db 67 VFTTKPKDIP--DYFKQSPFEGYSWERSMTFEDQGVCTVTSIDIKLEGDCPFYFIRYGVN 124
QY 130 FKEDGNILGHK-LEYNFISHNVYITADKQNGI-----KANFKIRHNI 171
Db 125 FPSSGPMOKKTLKWEPSFENMIV-----RDGVLIGDVSRTLLLEGDKHRCNFRSTYGA 179
QY 172 EDGSVOLADHYQONTPIGDPVLLPDNHYLSTQSALSQPKPKRDHVMVLEFVTAAGITH 231
Db 180 KKG-----VVLPEYFHV-----DHRII-----LSH 200
QY 232 GMD-----ELYK 238
Db 201 DKDYNTVEVYE 211

Search completed: April 19, 2004, 15:07:26
Job time : 46 secs

```

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF435427; AAL33912.1; --
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR000917; GFP_like.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP; 1.
 DR ProDom; PD013756; Green_fl_protein; 1.
 DR SEQUENCE 238 AA; 26997 MW; 5F80A192173CB84D CRC64;

Query Match 84.0%; Score 1062; DB 5; Length 238;
 Best Local Similarity 80.7%; Pred. No. 1.9e-83;
 Matches 192; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGNGVHGFVSVEGEGDGYKLTLCFTGKLPVWPPTL 60
 Db 1 MSKGEELFTGVVPIVLVDGNGVHGFVSVEGEGDGYKLTLCFTGKLPVWPPTL 60
 QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
 Db 61 VTTLGIGIQCFARYPEHMKNDFFKSAMPEGVVOERTIFFQDDGKYKTRGEVKFEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
 QY 181 HYQQTNPICDGPVLLPDNHYLSTQSALSADPKPKRDHMLVLEFVTAAGITHGMDELYK 238
 Db 181 HYQTNVPLGDGPVLLPINHYLSQTAISKDRNETRDHMLVLEFVTAAGITHGMDELYK 238

RESULT 11

Q8WTC8 PRELIMINARY; PRT; 238 AA.
 AC Q8WTC8;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Green fluorescent protein.
 GN GFP.

OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroids; Leptomedusae;
 OC Aequoreidae; Aequorea.
 OC NCBI_TaxID=147615;

RN [1]_TaxID=147615;
 RP SEQUENCE FROM N.A.
 RC STRAIN=GFPxm163;
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RA Li S.J., Xia N.S.;
 RT "Colorful mutants of green fluorescent protein from Aequorea
 macrodactyla.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435429; AAL33914.1; --
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR000917; GFP like.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP; 1.
 DR ProDom; PD013756; Green_fl_protein; 1.

DR SEQUENCE 238 AA; 27047 MW; 5F80A18FA1E7C84D CRC64;

Query Match 83.9%; Score 1061; DB 5; Length 238;
 Best Local Similarity 80.7%; Pred. No. 2.3e-83;
 Matches 192; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGNGVHGFVSVEGEGDGYKLTLCFTGKLPVWPPTL 60
 Db 1 MSKGEELFTGVVPIVLVDGNGVHGFVSVEGEGDGYKLTLCFTGKLPVWPPTL 60

QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
 Db 61 VTTLGIGIQCFARYPEHMKNDFFKSAMPEGVVOERTIFFQDDGKYKTRGEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
 QY 181 HYQQTNPICDGPVLLPDNHYLSTQSALSADPKPKRDHMLVLEFVTAAGITHGMDELYK 238
 Db 181 HYQTNVPLGDGPVLLPINHYLSQTAISKDRNETRDHMLVLEFVTAAGITHGMDELYK 238

RESULT 12

Q8WTC9 PRELIMINARY; PRT; 238 AA.
 AC Q8WTC9;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Green fluorescent protein.
 GN GFP.

OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroids; Leptomedusae;
 OC Aequoreidae; Aequorea.
 OC NCBI_TaxID=147615;

RN [1]_TaxID=147615;
 RP SEQUENCE FROM N.A.
 RC STRAIN=GFPxm162;
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RA Li S.J., Xia N.S.;
 RT "Colorful mutants of green fluorescent protein from Aequorea
 macrodactyla.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435428; AAL33913.1; --
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR000917; GFP like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.

DR PRINTS; PR01229; GFP; 1.
 DR ProDom; PD013756; Green_fl_protein; 1.
 DR SEQUENCE 238 AA; 27031 MW; 5F80A19C19DC584D CRC64;

Query Match 83.9%; Score 1061; DB 5; Length 238;
 Best Local Similarity 80.7%; Pred. No. 2.3e-83;
 Matches 192; Conservative 18; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGNGVHGFVSVEGEGDGYKLTLCFTGKLPVWPPTL 60
 Db 1 MSKGEELFTGVVPIVLVDGNGVHGFVSVEGEGDGYKLTLCFTGKLPVWPPTL 60

QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
 Db 61 VTTLGIGIQCFARYPEHMKNDFFKSAMPEGVVOERTIFFQDDGKYKTRGEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180

QY 181 HYQQTNPICDGPVLLPDNHYLSTQSALSADPKPKRDHMLVLEFVTAAGITHGMDELYK 238
 Db 181 HYQTNVPLGDGPVLLPINHYLSQTAISKDRNETRDHMLVLEFVTAAGITHGMDELYK 238

RESULT 13

Q8WTC7 PRELIMINARY; PRT; 238 AA.
 AC Q8WTC7;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Green fluorescent protein.
 GN GFP.

OS Aequorea macrodactyla.
 OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroids; Leptomedusae;
 OC Aequoreidae; Aequorea.
 OC NCBI_TaxID=147615;

RN [1]_TaxID=147615;
 RP SEQUENCE FROM N.A.
 RC STRAIN=GFPxm163;
 RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
 RA Li S.J., Xia N.S.;
 RT "Colorful mutants of green fluorescent protein from Aequorea
 macrodactyla.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF435429; AAL33914.1; --
 DR GO; GO:0006091; P:energy pathways; IEA.
 DR InterPro; IPR000917; GFP like.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP; 1.
 DR ProDom; PD013756; Green_fl_protein; 1.

DR SEQUENCE 238 AA; 27047 MW; 5F80A18FA1E7C84D CRC64;

```
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 27049 MW; 8185DOE5B529012B CRC64;

Query Match      85.2%; Score 1078; DB 5; Length 238;
Best Local Similarity 81.9%; Pred. No. 7.8e-85;
Matches 195; Conservative 18; Mismatches 25; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYGKLTGKLTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYGKLTGKLTCTTGKLPVWPPTL 60

QY 61 VTFPSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB 61 VTFPSYGIQCFARPPEHMKMNDFFKSAPEGYVOERTIFFKDDGNKYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGMDFKEDGNILGHKLEYNFIHSNVIYIMPDKANNGLKVNFKIRHNIEGGVQLAD 180

QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSQKPKERDHWMLLEFVTAAGITHGMDELYK 238
DB 181 HYQTNVPLGDGVPVLLPDNHYLSTQTAISKDRNETRDHMFLEFFSACGHTGMDLYK 238

RESULT 8
ID Q8WTC6 PRELIMINARY; PRT; 238 AA.
AC Q8WTC6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SG24;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RT "Colorful mutants of green fluorescent protein from Aequorea
macrodactyla.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435431; AAL33916.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 26956 MW; 75521EAP0CEBA73A CRC64;

Query Match      84.1%; Score 1064; DB 5; Length 238;
Best Local Similarity 81.9%; Pred. No. 1.2e-83;
Matches 195; Conservative 15; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYGKLTGKLTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYGKLTGKLTCTTGKLPVWPPTL 60

QY 61 VTFPSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB 61 VTFPSYGIQCFARPPEHMKMNDFFKSAPEGYVOERTIFFKDDGNKYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGMDFKEDGNILGHKLEYNFIHSNVIYIMPDKANNGLKVNFKIRHNIEGGVQLAD 180

QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSQKPKERDHWMLLEFVTAAGITHGMDELYK 238
DB 181 HYQTNVPLGDGVPVLLPDNHYLSTQTAISKDRNETRDHMFLEFFSACGHTGMDLYK 238

RESULT 10
ID Q8WTD0 PRELIMINARY; PRT; 238 AA.
AC Q8WTD0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GFPxnl61;
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Guan B.Q., Chen M.,
RA Li S.J., Xia N.S.;
RT "Colorful mutants of green fluorescent protein from Aequorea
macrodactyla.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF435431; AAL33916.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 688FD75B88926903 CRC64;

Query Match      84.7%; Score 1072; DB 5; Length 238;
Best Local Similarity 81.5%; Pred. No. 2.6e-84;
Matches 194; Conservative 18; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYGKLTGKLTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYGKLTGKLTCTTGKLPVWPPTL 60

QY 61 VTFPSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNKYKTRAEVKFEGDTLV 120
DB 61 VTFPSYGIQCFARPPEHMKMNDFFKSAPEGYVOERTIFFKDDGNKYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGMDFKEDGNILGHKLEYNFIHSNVIYIMPDKANNGLKVNFKIRHNIEGGVQLAD 180

QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALSQKPKERDHWMLLEFVTAAGITHGMDELYK 238
DB 181 HYQTNVPLGDGVPVLLPDNHYLSTQTAISKDRNETRDHMFLEFFSACGHTGMDLYK 238
```

```
Best Local Similarity 96.2%; Pred. No. 1.1e-97;
Matches 229; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYKGLTKLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYKGLTKLKFICTTGKLPVWPPTL 60
QY 61 VTFSGVQCFSRYPDHMKRHDFFKSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTFSGVQCFSRYPDHMKRHDFFKSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSQSALSQDPKPKEDHMLVLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSQSALSQDPKPKEDHMLVLEFVTAAGITHGMDELYK 238

RESULT 5
Q17105
ID Q17105 PRELIMINARY; PRT; 238 AA.
AC Q17105;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein (Fragment).
GN GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RP SEQUENCE FROM N.A.
RA Watkins J.N., Campbell A.K.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X83960; CAA58790.1; -
DR PIR; J06092; JQ1514.
DR HSSP; P42212; 1BFP.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPJORSSENT.
DR ProDom; PD013756; Green_fl_protein; 1.
FT NON_TER 238
SQ SEQUENCE 238 AA; 26950 MW; 26E2BE450E748E44 CRC64;

Query Match 94.2%; Score 1192; DB 5; Length 238;
Best Local Similarity 92.9%; Pred. No. 1.2e-94;
Matches 221; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYKGLTKLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYKGLTKLKFICTTGKLPVWPPTL 60
QY 61 VTFSGVQCFSRYPDHMKRHDFFKSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTFSGVQCFSRYPDHMKRHDFFKSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSQSALSQDPKPKEDHMLVLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSQSALSQDPKPKEDHMLVLEFVTAAGITHGMDELYK 238

RESULT 6
Q17106
ID Q17106 PRELIMINARY; PRT; 238 AA.

Best Local Similarity 96.2%; Pred. No. 1.1e-97;
Matches 229; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYKGLTKLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDGYKGLTKLKFICTTGKLPVWPPTL 60
QY 61 VTFSGVQCFSRYPDHMKRHDFFKSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTFSGVQCFSRYPDHMKRHDFFKSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVYITADKQNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSQSALSQDPKPKEDHMLVLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSQSALSQDPKPKEDHMLVLEFVTAAGITHGMDELYK 238

RESULT 7
Q8WP95
ID Q8WP95 PRELIMINARY; PRT; 238 AA.
AC Q8WP95;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
GN GFPXM.
OS Aequorea macrodactyla.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=147615;
RN [1]
RP SEQUENCE FROM N.A.
RA Luo W.X., Zhang J., Yang H.J., Xie X.Y., Li S.W., Qin Y.X., Pang S.Q.,
RA Li S.J., Xia N.S.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY013824; AAK02082.1; -
DR EMBL; AY013821; AAK02059.1; -
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPJORSSENT.
```

```

Db      61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY      121 NRLEKGDIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Db      121 NRLEKGDIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
QY      181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHVMVLEFVTAAGITHGMDELYK 238
Db      181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 2
Q8GHE4
ID      Q8GHE4      PRELIMINARY;      PRT;      238 AA.
AC      Q8GHE4
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Green fluorescence protein.
GN      375GFP.
OS      Azomonas agilis.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Azomonas.
CX      NCBI_TaxID=116849;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Koranyi P., Berenyi M., Burg K.;
RT      "Occurrence of green fluorescence protein in diazotrophic bacteria
RT      Azomonas and Azotobacter.";
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF324406; AAN86137.1; -.
DR      GO; GO:0006091; P:energy pathways; IEA.
DR      InterPro; IPR009017; GFP like.
DR      InterPro; IPR000786; Green_fl_protein.
DR      Pfam; PF01353; GFP; 1.
DR      PRINTS; PR01229; GFLUORESCENT.
DR      ProDom; PD013756; Green_fl_protein; 1.
DR      ProDom; PD013756; Green fl protein; 1.
SQ      SEQUENCE 238 AA; 26902 MW; 15FE9B9C5B4F6B89 CRC64;

Query Match      97.5%; Score 1233; DB 2; Length 238;
Best Local Similarity 97.1%; Pred. No. 3.5e-98;
Matches 231; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGDATYKLTIKFTCTGKLPVWPPTL 60
Db      1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGDATYKLTIKFTCTGKLPVWPPTL 60
QY      61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db      61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY      121 NRLEKGDIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Db      121 NRLEKGDIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
QY      181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHVMVLEFVTAAGITHGMDELYK 238
Db      181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 3
Q8GHE3
ID      Q8GHE3      PRELIMINARY;      PRT;      238 AA.
AC      Q8GHE3
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Green fluorescence protein.
GN      85GFP.
OS      Azotobacter vinelandii.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Azotobacter.
CX      NCBI_TaxID=354;

```

```

RN      [1]
RP      SEQUENCE FROM N.A.
RA      Koranyi P., Berenyi M., Burg K.;
RT      "Occurrence of green fluorescence protein in diazotrophic bacteria
RT      Azomonas and Azotobacter.";
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF324406; AAN86138.1; -.
DR      GO; GO:0006091; P:energy pathways; IEA.
DR      InterPro; IPR009017; GFP like.
DR      InterPro; IPR000786; Green_fl_protein.
DR      Pfam; PF01353; GFP; 1.
DR      PRINTS; PR01229; GFLUORESCENT.
DR      ProDom; PD013756; Green_fl_protein; 1.
DR      ProDom; PD013756; Green fl protein; 1.
SQ      SEQUENCE 238 AA; 26887 MW; E0E1616BD2AF6188 CRC64;

Query Match      97.4%; Score 1232; DB 2; Length 238;
Best Local Similarity 97.1%; Pred. No. 4.3e-98;
Matches 231; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGDATYKLTIKFTCTGKLPVWPPTL 60
Db      1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGDATYKLTIKFTCTGKLPVWPPTL 60
QY      61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db      61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY      121 NRLEKGDIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Db      121 NRLEKGDIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
QY      181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHVMVLEFVTAAGITHGMDELYK 238
Db      181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 4
Q93125
ID      Q93125      PRELIMINARY;      PRT;      238 AA.
AC      Q93125;
DT      01-FEB-1997 (TrEMBLrel. 02, Created)
DT      01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Green fluorescent protein mutant 3.
GN      GFP.
OS      Aequorea victoria (Jellyfish).
OC      Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroidea; Leptomedusae;
OC      Aequoreidae; Aequorea.
CX      NCBI_TaxID=6100;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Cormanck B.P., Valdivia R.H., Falkow S.;
RA      "FACS-optimized mutants of the green fluorescent protein (GFP).";
RL      Gene 173:33-38 (1996).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Cormanck B.P., Bertram G., Egerton M., Gow N.A.R., Falkow S.,
RA      Brown A.J.P.;
RT      "Yeast Enhanced Green Fluorescent Protein (yEGFP): a reporter of gene
RT      expression in Candida albicans.";
RL      Microbiology 0:0-0 (1996).
DR      EMBL; U73901; AAB18957.1; -.
DR      HSP; P42212; 1BFP.
DR      GO; GO:0006091; P:energy pathways; IEA.
DR      InterPro; IPR009017; GFP like.
DR      InterPro; IPR000786; Green_fl_protein.
DR      Pfam; PF01353; GFP; 1.
DR      PRINTS; PR01229; GFLUORESCENT.
DR      ProDom; PD013756; Green fl protein; 1.
DR      ProDom; PD013756; Green fl protein; 1.
SQ      SEQUENCE 238 AA; 26840 MW; A28622809A9D8A60 CRC64;

Query Match      97.0%; Score 1227; DB 5; Length 238;

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 15:02:40 ; Search time 45 Seconds
(without alignments)
1668.741 Million cell updates/sec

Title: US-10-057-505-2-COPY

Perfect score: 1265

Sequence: 1 MSKGEELFGVVPILVELDGLG.....VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1238	97.9	238	2 Q8GHE2	Q8ghe2 azotobacter
2	1233	97.5	238	2 Q8GHE4	Q8ghe4 azomonas ag
3	1232	97.4	238	2 Q8GHE3	Q8ghe3 azotobacter
4	1227	97.0	238	5 Q93125	Q93125 aequorea vi
5	1192	94.2	238	5 Q17105	Q17105 aequorea vi
6	1181	93.4	238	5 Q17106	Q17106 aequorea vi
7	1078	85.2	238	5 Q8WPC9	Q8wpc9 aequorea ma
8	1072	84.7	238	5 Q8WTC6	Q8wtc6 aequorea ma
9	1064	84.1	238	5 Q8WTC4	Q8wtc4 aequorea ma
10	1062	84.0	238	5 Q8WTD0	Q8wtd0 aequorea ma
11	1061	83.9	238	5 Q8WTC8	Q8wtc8 aequorea ma
12	1061	83.9	238	5 Q8WTC9	Q8wtc9 aequorea ma
13	1059	83.7	238	5 Q8WTC7	Q8wtc7 aequorea ma
14	1057	83.6	238	5 Q8WTC5	Q8wtc5 aequorea ma
15	247.5	19.6	225	5 Q95UA7	Q95ua7 montastraea
16	247.5	19.6	225	5 Q7Z0W5	Q7z0w5 montastraea

17	241.5	19.1	225	5	Q7Z0W9	Q7z0w9 montastraea
18	239.5	18.9	225	5	Q963F5	Q963f5 montastraea
19	236.5	18.7	225	5	Q816J8	Q816j8 trachyphyll
20	229	18.1	224	5	Q8MU48	Q8mu48 montastraea
21	228	18.0	225	5	Q8T5F1	Q8t5f1 montastraea
22	228	18.0	226	5	Q9U6I3	Q9u6i3 clavularia
23	225.5	17.8	225	5	Q7Z0W4	Q7z0w4 montastraea
24	224.5	17.7	236	5	Q8T6U0	Q8t6u0 dendronaph
25	208.5	16.5	259	5	Q8MMA2	Q8mma2 agaricia fr
26	208	16.4	227	5	Q7Z0W6	Q7z0w6 montastraea
27	207	16.4	239	5	Q8MMA1	Q8mma1 agaricia ag
28	206	16.3	229	5	Q9U6I6	Q9u6i6 anemonia ma
29	205.5	16.2	235	5	Q8T5F0	Q8t5f0 scolymia cu
30	203.5	16.1	214	5	Q86LV7	Q86lv7 meandrina m
31	202.5	16.0	214	5	Q86LV8	Q86lv8 meandrina m
32	202.5	16.0	234	5	Q7Z0W7	Q7z0w7 montastraea
33	202	16.0	234	5	Q8MU47	Q8mu47 montastraea
34	199.5	15.8	225	5	Q9U6Y8	Q9u6y8 discosoma s
35	199.5	15.8	227	5	Q962P9	Q962p9 montastraea
36	199.5	15.8	227	5	Q7Z0W8	Q7z0w8 montastraea
37	197.5	15.6	234	5	Q8T5F2	Q8t5f2 montastraea
38	196.5	15.5	232	5	Q9GP15	Q9gpi5 anemonia su
39	196	15.5	227	5	Q95VT0	Q95vt0 montastraea
40	195.5	15.5	225	5	Q8T6T9	Q8t6t9 radianthus
41	195.5	15.5	238	5	Q9BLY9	Q9bly9 renilla mue
42	194.5	15.4	232	5	Q9U6Y7	Q9u6y7 discosoma s
43	192.5	15.2	232	5	Q9GZ28	Q9gz28 anemonia su
44	190.5	15.1	222	5	Q7Z168	Q7z168 cerianthus
45	189	14.9	231	5	Q8T5F8	Q8t5f8 parasicyoni

ALIGNMENTS

RESULT 1
Q8GHE2 PRELIMINARY; PRT; 238 AA.
AC Q8GHE2
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Green fluorescence protein.
GN 2289GFP
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM2289;
RA Koranyi P., Berenyi M., Burg K.;
RT "Occurrence of green fluorescence protein in diazotrophic bacteria
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF324408; AAN86140.1; -
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP-like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 238 AA; 26914 MW; F84840F1F9064018 CRC64;

Query Match 97.9%; Score 1238; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 1.3e-98;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFGVVPILVELDGVNGHKFSVSGEGSDATYGLTKLFICTTKGLPVWPPTL 60
Db 1 MSKGEELFGVVPILVELDGVNGHKFSVSGEGSDATYGLTKLFICTTKGLPVWPPTL 60
QY 61 VTTSXGVQCFSRYPDHRHDFKFSAMPEGVVOERTFFKDDGNKYKTRAVKEGDPLV 120

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 15:04:50 ; Search time 22 seconds
(without alignments)
556.499 Million cell updates/sec

Title: US-10-057-505-2-COPY
Perfect score: 1265
Sequence: 1 MSKGEELFTGVVPIVELDGD.....VLLEFVTAAGTHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1238	97.9	238	1	US-08-753-143-2
2	1238	97.9	238	2	US-08-679-865-2
3	1238	97.9	238	2	US-08-680-876-2
4	1238	97.9	238	2	US-08-792-553-2
5	1238	97.9	238	3	US-08-753-144-2
6	1238	97.9	238	3	US-09-094-359-2
7	1238	97.9	238	3	US-09-172-063-2
8	1238	97.9	238	3	US-09-263-975-2
9	1238	97.9	238	4	US-08-727-452-2
10	1238	97.9	238	4	US-09-418-785-1
11	1238	97.9	238	4	US-09-129-192C-2
12	1238	97.9	238	4	US-09-129-192C-74
13	1238	97.9	238	4	US-09-602-641-2
14	1238	97.9	238	4	US-09-704-463-2
15	1237	97.8	247	3	US-08-893-327-20
16	1234	97.5	238	1	US-08-337-915A-2
17	1234	97.5	238	3	US-09-121-539-1
18	1234	97.5	238	4	US-09-214-909-2
19	1234	97.5	238	4	US-09-479-645A-10
20	1234	97.5	238	4	US-09-479-645A-159
21	1234	97.5	238	4	US-09-472-065A-4
22	1234	97.5	238	4	US-09-920-922-4
23	1234	97.5	238	5	PCT-US95-14692-2
24	1233	97.5	238	4	US-09-023-946B-35
25	1233	97.5	239	3	US-08-646-538-2
26	1233	97.5	239	3	US-09-503-222-2
27	1231	97.3	238	4	US-09-472-065A-2

28 1231 97.3 243 4 US-09-479-645A-94 Sequence 94, Appl
29 1231 97.3 243 4 US-09-479-645A-96 Sequence 96, Appl
30 1231 97.3 243 4 US-09-479-645A-98 Sequence 98, Appl
31 1231 97.3 243 4 US-09-479-645A-100 Sequence 100, Appl
32 1231 97.3 243 4 US-09-479-645A-102 Sequence 102, App
33 1231 97.3 243 4 US-09-479-645A-104 Sequence 104, App
34 1231 97.3 243 4 US-09-479-645A-110 Sequence 110, App
35 1231 97.3 1070 4 US-09-091-042A-2 Sequence 2, Appli
36 1231 97.3 1452 4 US-09-127-227-2 Sequence 2, Appli
37 1230 97.2 238 2 US-08-588-201-2 Sequence 2, Appli
38 1230 97.2 238 2 US-09-169-605-2 Sequence 2, Appli
39 1230 97.2 238 3 US-08-893-327-2 Sequence 2, Appli
40 1230 97.2 243 4 US-09-479-645A-88 Sequence 88, Appl
41 1230 97.2 243 4 US-09-479-645A-90 Sequence 90, Appl
42 1230 97.2 243 4 US-09-479-645A-92 Sequence 92, Appl
43 1230 97.2 243 4 US-09-479-645A-148 Sequence 148, App
44 1230 97.2 243 4 US-09-479-645A-150 Sequence 150, App
45 1230 97.2 243 4 US-09-479-645A-152 Sequence 152, App

ALIGNMENTS

RESULT 1
US-08-753-143-2
; Sequence 2, Application US/08753143A ✓ 577707A
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; FILE REFERENCE: 07257/032003
; CURRENT APPLICATION NUMBER: US/08/753.143A
; CURRENT FILING DATE: 1996-11-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-08-753-143-2

Query Match 97.9% Score 1238; DB 1; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MSKGEELFTGVVPIVELDGDVNGHKFSVSGEGSDATYKLTCLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIVELDGDVNGHKFSVSGEGSDATYKLTCLKFICTTGKLPVWPPTL 60
Qy 61 VTTFPSXGVQCFSRYPDHMKRHDFFKSAMPEGYVOERTIFFKDDGNYKTRAEVKEGDTLV 120
Db 61 VTTFPSXGVQCFSRYPDHMKRHDFFKSAMPEGYVOERTIFFKDDGNYKTRAEVKEGDTLV 120
Qy 121 NRLEKIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRLEKIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Qy 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKCRDHVLLFEVTAAGTHGMDELYK 238
Db 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKCRDHVLLFEVTAAGTHGMDELYK 238
RESULT 2
US-08-679-865-2
; Sequence 2, Application US/08679865
; Patent No. 5912137
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:

```
/
/ ADDRESS: Townsend and Townsend and Crew LLP
/ STREET: Two Embarcadero Center, Eighth Floor
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: USA
/ ZIP: 94111-3834
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/679,865
/ FILING DATE: 16-JUL-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Storella, John S.
/ REGISTRATION NUMBER: 32,944
/ REFERENCE/DOCKET NUMBER: 023072-069000
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 238 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-679-865-2

Query Match          97.9%; Score 1238; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIILVELDGDVNGHKFVSVEGEGDATYKGLTLKFICTTGKLPVWPPTL 60
Db 1 MSKGELFTGVVPIILVELDGDVNGHKFVSVEGEGDATYKGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGVVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGVVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGSGVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGSGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKEKDHMVLLFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKEKDHMVLLFVTAAGITHGMDELYK 238

RESULT 3
US-08-680-876-2
; Sequence 2, Application US/08680876
; Patent No. 5925558
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/680,876
/ FILING DATE: 16-JUL-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Storella, John S.
/ REGISTRATION NUMBER: 32,944
/ REFERENCE/DOCKET NUMBER: 023072-069200
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 238 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-680-876-2

Query Match          97.9%; Score 1238; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIILVELDGDVNGHKFVSVEGEGDATYKGLTLKFICTTGKLPVWPPTL 60
Db 1 MSKGELFTGVVPIILVELDGDVNGHKFVSVEGEGDATYKGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGVVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAPEGVVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGSGVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGSGVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKEKDHMVLLFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKEKDHMVLLFVTAAGITHGMDELYK 238

RESULT 4
US-08-792-553-2
; Sequence 2, Application US/08792553
; Patent No. 5981200
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Heim, Roger
; TITLE OF INVENTION: Tandem Fluorescent Protein Constructs
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSER: FISH & RICHARDSON P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/792,553
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lisa A. Haile, Ph.D.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/041001/UC 96-160-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 2:
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-792-553-2

Query Match          97.9%; Score 1238; DB 2; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKLTLLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKLTLLKFICTTGKLPVWPPTL 60
QY 61 VTTFSXGVQCFSRYPDHHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSXGVQCFSRYPDHHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVTITADKQKNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVTITADKQKNGIKANFKIRHNIEDGSVOLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHMLLEFVTAAGITHGMDELYK 238

```

```

RESULT 5
US-08-753-144-2
; Sequence 2, Application US/08753144
; Patent No. 6066476
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Heim, Roger
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/753,144
; FILING DATE: 20-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/727,452
; FILING DATE: 10-OCT-1996
; APPLICATION NUMBER: US95/14692
; FILING DATE: 13-NOV-1995
; APPLICATION NUMBER: 08/337,915
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/032002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

```

```

US-08-753-144-2

Query Match          97.9%; Score 1238; DB 3; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKLTLLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKLTLLKFICTTGKLPVWPPTL 60
QY 61 VTTFSXGVQCFSRYPDHHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSXGVQCFSRYPDHHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVTITADKQKNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVTITADKQKNGIKANFKIRHNIEDGSVOLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 6
US-09-094-359-2
; Sequence 2, Application US/09094359
; Patent No. 6140132
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/067001
; CURRENT APPLICATION NUMBER: US/09/094,359
; CURRENT FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-094-359-2

Query Match          97.9%; Score 1238; DB 3; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKLTLLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKLTLLKFICTTGKLPVWPPTL 60
QY 61 VTTFSXGVQCFSRYPDHHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSXGVQCFSRYPDHHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVTITADKQKNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNVTITADKQKNGIKANFKIRHNIEDGSVOLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 7
US-09-172-063-2
; Sequence 2, Application US/09172063
; Patent No. 6150176
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan

```

```

; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; EARLIER FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
; EARLIER FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-172-063-2

Query Match          97.9%; Score 1238; DB 3; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGKLTLCFTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGKLTLCFTCTTGKLPVWPPTL 60

QY 61 VTTSXGVQCFSPYDPDMKRDHDFPKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTSYGVQCFSPYDPDMKRDHDFPKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180

QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKRDHVMVLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 8
US-09-263-975-2
; Sequence 2, Application US/09263975
; Patent No. 6248550
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Assays for Protein Kinases Using
; TITLE OF INVENTION: Fluorescent Protein Substrates
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,975
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/679,865
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Storella, John S.
; REGISTRATION NUMBER: 32,944
; REFERENCE/DOCKET NUMBER: 02307Z-069000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200

```

```

; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-263-975-2

Query Match          97.9%; Score 1238; DB 3; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGKLTLCFTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGKLTLCFTCTTGKLPVWPPTL 60

QY 61 VTTSXGVQCFSPYDPDMKRDHDFPKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTSYGVQCFSPYDPDMKRDHDFPKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180

QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKRDHVMVLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 9
US-08-727-452-2
; Sequence 2, Application US/08727452A
; Patent No. 6319669
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Heim, Roger
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR DETECTION OF ANALYTES
; FILE REFERENCE: 07257/032001
; CURRENT APPLICATION NUMBER: US/08/727,452A
; EARLIER FILING DATE: 1996-03-20
; EARLIER APPLICATION NUMBER: PCT/US95/14692
; EARLIER FILING DATE: 1995-11-13
; EARLIER APPLICATION NUMBER: US 07/337,915
; EARLIER FILING DATE: 1994-11-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-08-727-452-2

Query Match          97.9%; Score 1238; DB 4; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGKLTLCFTCTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGKLTLCFTCTTGKLPVWPPTL 60

QY 61 VTTSXGVQCFSPYDPDMKRDHDFPKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTSYGVQCFSPYDPDMKRDHDFPKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120

QY 121 NRIELKGIDFKEDGNILGHKLEYNFIISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180

QY 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKRDHVMVLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDPVLLPDNHYLSTQSALSKDPKRDHVMVLEFVTAAGITHGMDELYK 238

```

```
RESULT 10
US-09-418-785-1
; Sequence 1, Application US/09418785
; Patent No. 6414119
; GENERAL INFORMATION:
; APPLICANT: Fisher, Hugh
; TITLE OF INVENTION: Rapidly Greening, Low Oxygen Mutant of
; TITLE OF INVENTION: the Aequorea victoria Green Fluorescent Protein
; FILE REFERENCE: RUC 99-0011
; CURRENT APPLICATION NUMBER: US/09/418,785
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/104,563
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
; PUBLICATION INFORMATION:
; AUTHORS: Prasher, D.C. et al.
; TITLE: Primary structure of the Aequorea victoria green-f
; JOURNAL: Gene
; VOLUME: 111
; PAGES: 229-233
; DATE: 1992-01-01
; DATABASE ENTRY NUMBER: Genbank M62653
; DATABASE ACCESSION DATE: 1993-04-26
US-09-418-785-1

Query Match          97.9%; Score 1238; DB 4; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIITADKQKNGIKANFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIITADKQKNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 11
US-09-129-192C-2
; Sequence 2, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea
US-09-129-192C-2

Query Match          97.9%; Score 1238; DB 4; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIITADKQKNGIKANFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIITADKQKNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 12
US-09-129-192C-74
; Sequence 74, Application US/09129192C
; Patent No. 6495664
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/09/129,192C
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-09-129-192C-74

Query Match          97.9%; Score 1238; DB 4; Length 238;
Best Local Similarity 97.5%; Pred. No. 6.8e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
DB 1 MSKGEELFTGVVPIVLVDGVDVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
DB 61 VTTFSYGVQCFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIITADKQKNGIKANFKIRHNIEDGSVQLAD 180
DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVIITADKQKNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 13
US-09-602-641-2
; Sequence 2, Application US/09602641
; Patent No. 6608189
; GENERAL INFORMATION:
; APPLICANT: Taisen, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wachter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/602,641
```

```

; CURRENT FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/172,063
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-602-641-2

Query Match
Best Local Similarity 97.9%; Score 1238; DB 4; Length 238;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVPWPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKFSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKFSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTTAAGITHGMDELYK 238

RESULT 14
US-09-704-463-2
; Sequence 2, Application US/09704463
; Patent No. 6627449
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: 07257/067001
; CURRENT APPLICATION NUMBER: US/09/704,463
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 09/094,359
; PRIOR FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-704-463-2

Query Match
Best Local Similarity 97.9%; Score 1238; DB 4; Length 238;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVPWPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKFSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKFSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTTAAGITHGMDELYK 238
```

```

Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTTAAGITHGMDELYK 238

RESULT 15
US-08-893-327-20
; Sequence 20, Application US/08893327
; Patent No. 6020192
; GENERAL INFORMATION:
; APPLICANT: Zolotukhin, Sergei
; APPLICANT: Hauswirth, William W.
; APPLICANT: Muzyczka, Nicholas
; TITLE OF INVENTION: Humanized Green Fluorescent Protein
; TITLE OF INVENTION: Genes and Methods
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,327
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/588,201
; FILING DATE: 18-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UFLA:062\KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-893-327-20

Query Match 97.8%; Score 1237; DB 3; Length 247;
Best Local Similarity 97.5%; Pred. No. 9.3e-127;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVPWPTL 60
Db 10 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYGLTLKFICTTGKLPVPWPTL 69
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKFSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 70 VTTFSGVQCFSRYPDHMKRHDFFKFSAMPEGVYQERTIFFKDDGNYKTRAEVKFEGDTLV 129
QY 121 NRIELKGIDFEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVQLAD 180
Db 130 NRIELKGIDFEDGNILGHKLEYNFIHSNVIYITADKQNGIKANFKIRHNIEDGSVQLAD 189
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTTAAGITHGMDELYK 238
Db 190 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLLEFVTTAAGITHGMDELYK 247
```

Search completed: April 19, 2004, 15:08:34
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 14:55:50 ; Search time 60 Seconds
(without alignments)
1120.772 Million cell updates/sec

Title: US-10-057-505-2-COPY

Perfect score: 1265

Sequence: 1 MSKGEELFTGVVPILVELDGL.....VLLEFVTAAGTGHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1259	99.5	238	2 AAW05309	Aaw05309 Green flu
2	1259	99.5	238	6 ABG76006	Abg76006 Jellyfish
3	1257	99.4	501	2 AAW31878	Aaw31878 GFP varia
4	1257	99.4	501	2 AAW31879	Aaw31879 GFP varia
5	1250	98.8	238	6 ABG76009	Abg76009 Jellyfish
6	1242	98.2	238	4 AAB73554	Aab73554 Green flu
7	1242	98.2	238	6 ADA25213	Ada25213 Green flu
8	1241	98.1	238	2 AAW05310	Aaw05310 Green flu
9	1241	98.1	238	6 ABG76005	Abg76005 Jellyfish
10	1241	98.1	238	6 AAE34991	Aae34991 Aequorea
11	1241	98.1	501	2 AAW31880	Aaw31880 GFP varia
12	1241	98.1	514	2 AAW31876	Aaw31876 GFP varia
13	1240	98.0	238	6 AAE34992	Aae34992 Aequorea
14	1240	98.0	238	6 AAE34993	Aae34993 Aequorea
15	1239	97.9	238	2 AAW05311	Aaw05311 GFP mutan
16	1239	97.9	238	6 AAE34998	Aae34998 Aequorea
17	1238	97.9	238	2 AAW05304	Aaw05304 Green flu
18	1238	97.9	238	2 AAW05308	Aaw05308 Green flu
19	1238	97.9	238	2 AAW24232	Aaw24232 Aequorea
20	1238	97.9	238	2 AAW76105	Aaw76105 A. victor
21	1238	97.9	238	2 AAW40479	Aaw40479 A. victor
22	1238	97.9	238	2 AAW65081	Aaw65081 A. victor
23	1238	97.9	238	4 AAW76371	Aaw76371 A. victor
24	1238	97.9	238	4 AAB73552	Aab73552 Wild-type
25	1238	97.9	238	5 AAE16055	Aae16055 Aequorea

26	1238	97.9	238	5 AAE16056	Aae16056 Aequorea
27	1238	97.9	238	5 AAE16038	Aae16038 Aequorea
28	1238	97.9	238	5 ABG32365	Abg32365 Aequorea
29	1238	97.9	238	6 ABG76012	Abg76012 Jellyfish
30	1238	97.9	238	6 ABG75980	Abg75980 Jellyfish
31	1238	97.9	238	6 AAE34999	Aae34999 Aequorea
32	1238	97.9	238	6 AAE34993	Aae34993 Aequorea
33	1238	97.9	238	6 ABR44423	Abra44423 Wild-type
34	1238	97.9	238	6 ADA25219	Ada25219 Green flu
35	1238	97.9	238	6 ADA25194	Ada25194 Aequorea
36	1238	97.9	501	2 AAW31877	Aaw31877 GFP varia
37	1237	97.8	238	2 AAW22100	Aaw22100 Aequorea
38	1237	97.8	238	2 AAW65084	Aaw65084 A. victor
39	1237	97.8	238	2 AAW65082	Aaw65082 A. victor
40	1237	97.8	238	2 AAY43248	Aay43248 Green flu
41	1237	97.8	238	4 AAB73555	Aab73555 Green flu
42	1237	97.8	238	5 AAE16057	Aae16057 Aequorea
43	1237	97.8	238	6 ABG76013	Abg76013 Jellyfish
44	1237	97.8	238	6 AAE34990	Aae34990 Aequorea
45	1237	97.8	247	2 AAW96330	Aaw96330 Humanised

ALIGNMENTS

RESULT 1
AAW05309
ID AAW05309 standard; protein; 238 AA.
XX AC AAW05309;
XX
DT 02-APR-1997 (first entry)
XX
DE Green fluorescent protein mutant Y66W/N146I/M153T/V163A/N212K.
XX
KW Green fluorescent protein; GFP; jellyfish; Aequorea victoria; sea pansy;
KW Renilla reniformis; differential gene expression; protein localisation;
KW gene expression tracking; fluorescence.
XX OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 66 /note= "Y66W"
FT Misc-difference 146 /note= "N146I"
FT Misc-difference 153 /note= "M153T"
FT Misc-difference 163 /note= "M153T"
FT Misc-difference 212 /note= "V163A"
FT Misc-difference 212 /note= "N212K"
W09623810-A1.
PD ~~08-AUG-1996~~
XX
PF 13-NOV-1995; 95WO-US014692.
XX
PR 10-NOV-1994; 94US-00337915.
PA (REGC) UNIV CALIFORNIA.
XX
PI Tsien RY, Heim R;
XX
DR WPI; 1996-371370/37.
XX
PT New modified Aequorea green fluorescent polypeptide(s) - having amino
PT acid changes to provide prods. which exhibit different excitation and
XX emission spectra.
XX
PS Claim 11; Page ?; 39pp; English.
XX

XX GFP variants S65C and W7 tandem fluorescent protein construct.

XX North West Pacific jellyfish; green fluorescent protein; GFP; S65C;

XX fluorescent resonance energy transfer; FRET; enzymatic assay; W7;

XX enzyme inhibitor; enzyme promoter; PCR primer; protease cleavage site;

XX Tandem fluorescent protein construct; blue fluorescent protein.

XX Synthetic.

OS Aequorea victoria.

XX Key Location/Qualifiers

FT Protein 1..238

FT /label= S65C_GFP_variant

FT Misc-difference 65

FT /label= S65C

FT /note= "wild type Ser substituted with Cys"

FT Peptide 239..264

FT /label= linker_moeity

FT Cleavage-site 250..251

FT /label= trypsin_cleavage_site

FT Cleavage-site 253..254

FT /label= calpain_cleavage_site

FT Cleavage-site 258..259

FT /label= trypsin_enterokinase_cleavage_site

FT Protein 265..501

FT /label= W7_GFP_variant

FT Misc-difference 329

FT /label= Y66W

FT /note= "wild type Tyr substituted with Trp"

FT Misc-difference 409

FT /label= N146L

FT /note= "wild type Asn substituted with Leu"

FT Misc-difference 416

FT /label= M153T

FT /note= "wild type Met substituted with Thr"

FT Misc-difference 426

FT /label= V163A

FT /note= "wild type Val substituted with Ala"

FT Misc-difference 475

FT /label= N212K

FT /note= "wild type Asn substituted with Lys"

XX WO9728261-A1.

XX 07-AUG-1997.

XX 31-JAN-1997; 97WO-US001457.

XX 31-JAN-1996; 96US-00594575.

XX (REGC) UNIV CALIFORNIA.

XX (AURO-) AURORA BIOSCIENCES CORP.

XX Tsien RY, Heim R, Cubitt A;

XX WPI; 1997-402615/37.

XX Tandem fluorescent protein constructs - have donor and acceptor moieties

XX exhibiting fluorescent linked via cleavable peptide linker, useful in

XX enzymatic assays.

XX Claim 3; Page; 88pp; English.

XX This protein sequence is that of a novel tandem fluorescent protein

XX construct, made using Aequorea victoria (North West Pacific jellyfish)

XX green fluorescent protein (GFP) variants S65C and W7. W7 fluoresces at a

XX shorter wavelength than GFP. The construct comprises a donor (e.g. S65C)

XX and an acceptor (e.g. W7) fluorescent protein moiety (donors and

XX acceptors can be green or blue fluorescent proteins), and a linker

XX coupling them. Preferably, the donor is positioned at the N-terminus of

XX the polypeptide relative to the acceptor. The linker moiety is a peptide

XX 5-50 amino acids in length containing a protease cleavage site. In this

CC example, the linker moiety contains many recognition sites for proteases,

CC including trypsin, calpain and enterokinase. The donor and acceptor

CC moieties exhibit fluorescent resonance energy transfer (FRET) when they

CC are cleaved. The constructs are used in enzymatic assays and can be used

CC to isolate new enzymes or enzyme inhibitors or promoters. The specific

CC activity of enzyme (in vivo and in vitro) and compounds altering enzyme

CC activity can be obtained. FRET and hence activity of specific compounds

CC is measured from the acceptor or donor moiety or maybe obtained using a

CC ratio between the two. Note: The present sequence does not appear in the

CC specification; it has been made by modifying the native GFP sequence, and

CC adding the linker moiety in the appropriate place

XX Sequence 501 AA;

Query Match 99.4%; Score 1257; DB 2; Length 501;

Best Local Similarity 98.7%; Pred. No. 8e-122;

Matches 235; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGDATYKLTLCFTCTGKLPVWPTL 60

Db 264 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGDATYKLTLCFTCTGKLPVWPTL 323

QY 61 VTFPSXGVQCFSRYPDHMKRHDFFKSNMPGYVQERTIFFKDDGNYKTRAVKPEGDTLV 120

Db 324 VTFPSXGVQCFSRYPDHMKRHDFFKSNMPGYVQERTIFFKDDGNYKTRAVKPEGDTLV 383

QY 121 NRLEKGIIDFKEDGNILGHKLEYNFISHNVIITADKQNGIKANFKIRHNIEDGSVOLAD 180

Db 384 NRLEKGIIDFKEDGNILGHKLEYNFISHNVIITADKQNGIKANFKIRHNIEDGSVOLAD 443

QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKPKRKHMHVLEFVTAAGITHGMDLYK 238

Db 444 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKPKRKHMHVLEFVTAAGITHGMDLYK 501

RESULT 4

AAW31879

ID AAW31879 standard; protein; 501 AA.

XX AAW31879;

XX 03-FEB-1998 (first entry)

XX GFP variants S65T and W7 tandem fluorescent protein construct.

XX North West Pacific jellyfish; green fluorescent protein; GFP; S65C;

XX fluorescent resonance energy transfer; FRET; enzymatic assay; W7;

XX enzyme inhibitor; enzyme promoter; PCR primer; protease cleavage site;

XX Tandem fluorescent protein construct; blue fluorescent protein.

XX Synthetic.

OS Aequorea victoria.

XX Key Location/Qualifiers

FT Protein 1..238

FT /label= S65T_GFP_variant

FT Misc-difference 65

FT /label= S65C

FT /note= "wild type Ser substituted with Thr"

FT Peptide 239..264

FT /label= linker_moeity

FT Cleavage-site 250..251

FT /label= trypsin_cleavage_site

FT Cleavage-site 253..254

FT /label= calpain_cleavage_site

FT Cleavage-site 258..259

FT /label= trypsin_enterokinase_cleavage_site

FT Protein 265..501

FT /label= W7_GFP_variant

FT Misc-difference 329

FT /label= Y66W

FT /note= "wild type Tyr substituted with Trp"

FT Misc-difference 409

FT /label= N146L
 FT /note= "wild type Asn substituted with Leu"
 FT Misc-difference 416
 FT /label= M153T
 FT /note= "wild type Met substituted with Thr"
 FT Misc-difference 426
 FT /label= V163A
 FT /note= "wild type Val substituted with Ala"
 FT Misc-difference 475
 FT /label= N212K
 FT /note= "wild type Asn substituted with Lys"
 FT
 PN W09728261-Al.
 XX
 XX 07-AUG-1997.
 PD
 PD 31-JAN-1997; 97WO-US001457.
 PF
 XX 31-JAN-1996; 96US-00594575.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA (AURO-) AURORA BIOSCIENCES CORP.
 XX
 XX Tsien RY, Heim R, Cubitt A;
 PI WPI; 1997-402615/37.
 DR
 XX

Tandem fluorescent protein constructs - have donor and acceptor moieties exhibiting fluorescent linked via cleavable peptide linker, useful in enzymatic assays.

Claim 3; Page; 88pp; English.

This protein sequence is that of a novel tandem fluorescent protein construct, made using Aequorea victoria (North West Pacific Jellyfish) green fluorescent protein (GFP) variants S65T and W7. W7 fluoresces at a shorter wavelength than GFP. The construct comprises a donor (e.g. S65T) and an acceptor (e.g. W7) fluorescent protein moieties (donors and acceptors can be green or blue fluorescent proteins), and a linker coupling them. Preferably, the donor is positioned at the N-terminus of the polypeptide relative to the acceptor. The linker moiety is a peptide 5-50 amino acids in length containing a protease cleavage site. In this example, the linker moiety contains many recognition sites for proteases, including trypsin, calpain and enterokinase. The donor and acceptor moieties exhibit fluorescent resonance energy transfer (FRET) when they are cleaved. The constructs are used in enzymatic assays and can be used to isolate new enzymes or enzyme inhibitors or promoters. The specific activity of enzyme (in vivo and in vitro) and compounds altering enzyme activity can be obtained. FRET and hence activity of specific compounds is measured from the acceptor or donor moiety or maybe obtained using a ratio between the two. Note: The present sequence does not appear in the specification; it has been made by modifying the native GFP sequence, and adding the linker moiety in the appropriate place

Sequence 501 AA;

Query Match 99.4%; Score 1257; DB 2; Length 501;
 Best Local Similarity 98.7%; Pred. No. 8e-122;
 Matches 235; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPLVILVDGVNKHFSVSGEGDATTYKLTUKFTCTGKLPVMPPTL 60
 Db 264 MSKGEELFTGVVPLVILVDGVNKHFSVSGEGDATTYKLTUKFTCTGKLPVMPPTL 323
 QY 61 VTFPSXGVCQFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNKYTRAEVKEGTLV 120
 Db 324 VTFPSGWGVCQFSRYPDHMKRHDFFKFSAMPEGYVQERTIFFKDDGNKYTRAEVKEGTLV 383
 QY 121 NRIELKGIDFKEDGNILGHKLEYNFI SHNVYITADKQNGIKANFKIRHNIEDGSVOLAD 180
 Db 384 NRIELKGIDFKEDGNILGHKLEYNLSHVNVITADKQNGIKANFKIRHNIEDGSVOLAD 443
 QY 181 HYQONTPIGDGPVLLPNHLYLSTQSALS KDPKPKRDMHVLLEFVTAAGITHGMDELYK 238

Db 444 HYQONTPIGDGPVLLPNHLYLSTQSALS KDPKPKRDMHVLLEFVTAAGITHGMDELYK 501
 RESULT 5
 ABG76009
 ID ABG76009 standard; protein; 238 AA.
 XX AC ABG76009;
 XX 30-APR-2003 (first entry)
 XX Jellyfish GFP mutant W1B.
 XX Jellyfish; enzyme; green fluorescent protein; GFP; FRET; mutant;
 KW fluorescent resonance energy transfer; tandem fluorescent protein;
 KW enzymatic assay; Alzheimer's disease; hypertension; inflammation;
 KW apoptosis; AIDS; acquired immunodeficiency syndrome; mitein.
 XX Aequorea victoria.
 OS Synthetic.

Key Location/Qualifiers
 FT Misc-difference 64 /note= "Wild-type Phe substituted by Leu"
 FT Misc-difference 65 /note= "Wild-type Ser substituted by Thr"
 FT Misc-difference 66 /note= "Wild-type Tyr substituted by Trp"
 FT Misc-difference 146 /note= "Wild-type Asn substituted by Ile"
 FT Misc-difference 153 /note= "Wild-type Met substituted by Thr"
 FT Misc-difference 163 /note= "Wild-type Val substituted by Ala"
 FT Misc-difference 212 /note= "Wild-type Asn substituted by Lys"
 US2002164674-A1.
 07-NOV-2002.
 25-JAN-2002; 2002US-00057505.
 31-JAN-1996; 96US-00594575.
 31-JAN-1997; 97US-00792553.
 13-SEP-1999; 99US-00396003.
 (REGC) UNIV CALIFORNIA.
 Tsien RY, Heim R, Cubitt A;
 WPI; 2003-247255/24.

New tandem fluorescent protein construct comprising a donor or acceptor fluorescent protein moiety or a linker moiety that couples the donor and acceptor moieties, useful in enzymatic assays.

Claim 18; Page; 34pp; English.

The invention relates to a tandem fluorescent protein construct comprising a donor or acceptor fluorescent protein moiety (e.g. jellyfish green fluorescent protein, GFP) or a linker moiety that couples the donor and acceptor moieties. Also include are a recombinant nucleic acid coding for expression of the tandem fluorescent protein construct, an expression vector comprising expression control sequences operatively linked to a sequence coding for the expression of the tandem fluorescent protein construct, a host cell transfected with the expression vector, determining whether a sample contains an enzyme or whether a compound alters the activity of an enzyme, determining the amount of activity of an enzyme in a cell and testing for cleavage enzyme activity. The tandem fluorescent protein construct is useful in enzymatic assays, using the principle of fluorescent resonance energy transfer (FRET) between the

CC donor and acceptor moieties. The tandem fluorescent proteins are
 CC particularly useful in assays for protease activity. Proteases play an
 CC essential role in many disease processes e.g. Alzheimer's disease,
 CC hypertension, inflammation, apoptosis and AIDS (acquired immunodeficiency
 CC syndrome). Tandem fluorescent proteins were constructed comprising
 CC mutants of GFP with altered fluorescent spectra. The present sequence
 CC represents a mutant GFP suitable for incorporation into a tandem
 CC fluorescent protein of the invention. Note: The present sequence is not
 CC shown in the specification but was created by the indexer using the
 CC information in the specification and the wild-type GFP protein appearing
 CC as ABG75980
 XX
 XX Sequence 238 AA;

Query Match 98.8%; Score 1250; DB 6; Length 238;
 Best Local Similarity 98.3%; Pred. No. 1.4e-121;
 Matches 234; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDGYGKLTGKLTGKLPVWPPTL 60
 DB 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDGYGKLTGKLTGKLPVWPPTL 60
 QY 61 VTTSXGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKPEGDTLV 120
 DB 61 VTTLTWGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKPEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVITADKQNGIKANFKIRHNIEDGSVQLAD 180
 DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVITADKQNGIKANFKIRHNIEDGSVQLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKEKRDHMLVLEFVTAAGITHGMDELYK 238
 DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKEKRDHMLVLEFVTAAGITHGMDELYK 238

RESULT 6
 AAB73554
 ID AAB73554 standard; protein; 238 AA.
 XX
 AC AAB73554;
 XX
 XX 07-AUG-2001 (first entry)
 XX
 DE Green fluorescent protein mutant, GFPRI.
 XX
 KW Phenotype selection; non-selectable; fusion protein; stable expression;
 KW selectable marker; antibiotic resistance gene; Escherichia coli;
 KW green fluorescent protein; GFP; GFPRI; pGFP; mutant; mitein.
 XX
 OS Aequorea victoria.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX Misc-difference 163 /note= "Ala replaces wild-type Val"
 XX
 XX WO200129225-A1.
 XX
 XX 26-APR-2001.
 XX
 XX 29-MAR-2000; 2000WO-US008477.
 XX
 XX 21-OCT-1999; 99US-0160461P.
 XX 22-FEB-2000; 2000US-00510097.
 XX
 XX (PANO-) PANORAMA RES INC.
 XX
 XX Balint RF;
 XX
 XX WPI; 2001-282162/29.
 XX N-PSDB; AAH20247.
 XX
 XX Obtaining cells expressing mutant protein, comprises selecting from cells

PT transformed with library of mutagenized protein coding sequences joined
 PT to selector protein, which confers growth under selective conditions.
 XX
 PS Example 2; Page; 52pp; English.

CC The invention relates to methods whereby variants of proteins which do
 CC not confer selectable phenotypes can be selected for stable expression in
 CC host cells (especially Escherichia coli). The methods can be used to
 CC obtain mutants of a desired protein optimised for expression in the host
 CC cells, for obtaining a mutant protein with enhanced stability relative to
 CC the wild-type protein, and for identifying peptides that stabilise an
 CC unstable protein. The methods all involve expressing the protein of
 CC interest as a fusion with a protein that can confer a selectable
 CC phenotype, such as an antibiotic resistance protein. The transformed host
 CC cells are then grown under selective pressure (e.g., presence of
 CC antibiotic). The cells able to grow under such conditions are those which
 CC contain fusion proteins which are optimised for expression or which are
 CC more stable, as this property will also correlate with an increased
 CC amount or rate of synthesis of the selectable marker. The invention also
 CC discloses mutants of green fluorescent protein (GFP) selected for
 CC increased stability using the method of the invention. The present
 CC sequence represents a GFP mutant (GFPRI) which has increased stability
 CC relative to wild-type GFP. The present sequence is not shown in the
 CC specification, but was derived from the plasmid pGFP (GenBank accession
 CC number U17997) and the information given on page 23 of the specification
 XX
 XX Sequence 238 AA;

Query Match 98.2%; Score 1242; DB 4; Length 238;
 Best Local Similarity 97.9%; Pred. No. 9.8e-121;
 Matches 233; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDGYGKLTGKLTGKLPVWPPTL 60
 DB 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDGYGKLTGKLTGKLPVWPPTL 60
 QY 61 VTTSXGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKPEGDTLV 120
 DB 61 VTTSYGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKPEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVITADKQNGIKANFKIRHNIEDGSVOLAD 180
 DB 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVITADKQNGIKANFKIRHNIEDGSVOLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKEKRDHMLVLEFVTAAGITHGMDELYK 238
 DB 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMLVLEFVTAAGITHGMDELYK 238

RESULT 7
 ADA25213
 ID ADA25213 standard; protein; 238 AA.
 XX
 AC ADA25213;
 XX
 XX 20-NOV-2003 (first entry)
 XX
 DE Green fluorescent protein mutant, VI63A-GFP.
 XX
 KW Green fluorescent protein; GFP; jellyfish; marker protein;
 KW reporter protein; mutant; mitein.
 XX
 OS Synthetic.
 OS Aequorea victoria.
 XX
 XX Key Location/Qualifiers
 XX Misc-difference 163 /note= "Ala replaces wild-type Val"
 XX
 XX GB2374868-A.
 XX
 XX 30-OCT-2002.
 XX

28-SEP-2001; 2001GB-00023288.
 23-APR-2001; 2001GB-00009858.
 (AMSH) AMERSHAM PHARMACIA BIOTECH UK LTD.
 (AMSH) AMERSHAM BIOSCIENCES UK LTD.
 Stubbs SLJ, Jones AE, Michael NP, Thomas N;
 WPI; 2003-095652/09.
 N-PSDB; ADA25193.
 Novel fluorescent protein derived from green fluorescent protein useful
 as a transfection marker, has different excitation spectrum and/or
 emission spectrum compared with wild-type green fluorescent protein.
 Example 2; Page; 52pp; English.
 The invention relates to Aequorea victoria green fluorescent protein
 (GFP) mutants containing an amino acid substitution at positions 64 and
 175, and additionally an amino acid substitution at either position 65
 or 222. The mutants of the invention are particularly F64L-S175G-E222G-
 GFP (ADA25195) or F64L-S65T-S175G-GFP (ADA25196). GFP mutants of the
 invention exhibit enhanced fluorescence relative to wild type GFP when
 expressed in non-homologous cells at temperatures above 30 degrees
 Celsius, and excited at 490 nm. The mutants can also be detected in
 mammalian cells at lower levels of expression and with increased
 sensitivity relative to wild type GFP. The GFP mutants of the invention
 are useful as non-toxic markers for selection of transfected cells, as
 protein labels in living and fixed cells, as markers in cell or organelle
 fusion, for visualising translocation of intracellular proteins to a
 specific organelle, as secretion markers, as genetic reporters or protein
 tags for protein and gene expression in transgenic animals, as cell or
 organelle integrity markers, as transfection markers, as markers to be
 used in combination with fluorescent activated cell sorting (FACS), as
 real-time probes working at near physiological concentrations, for
 performing transposon vector mutagenesis, and as reporters for bacterial
 detection. The present sequence represents an Aequorea victoria GFP
 mutant used in an example of the invention. Note: The present sequence is
 not shown in the specification, but is derived from the wild-type GFP
 sequence shown in Fig 2 and the information given on page 24.
 Sequence 238 AA;
 Query Match 98.2%; Score 1242; DB 6; Length 238;
 Best Local Similarity 97.9%; Pred. No. 9, 8e-121;
 Matches 233; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPIVLVDGVDGNGHFKFSVSGEGDATYGKLTILKFICTTGKLPVPPPTL 60
 Db 1 MSKGEELFTGVVPIVLVDGVDGNGHFKFSVSGEGDATYGKLTILKFICTTGKLPVPPPTL 60
 QY 61 VTFSSGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLV 120
 Db 61 VTFSSGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLV 120
 QY 121 NRLEKIGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 Db 121 NRLEKIGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALS KDPKPKRDHVMVLEFVTAAGITHGMDELYK 238
 Db 181 HYQONTPIGDGVPVLLPDNHYLSTQSALS KDPKPKRDHVMVLEFVTAAGITHGMDELYK 238
 RESULT 8
 AAW05310
 ID AAW05310 standard; protein; 238 AA.
 XX
 AC AAW05310;
 XX
 DT 02-APR-1997 (first entry)
 XX

DE Green fluorescent protein mutant Y66H/Y145F.
 XX
 KW Green fluorescent protein; GFP; jellyfish; Aequorea victoria; sea pansy;
 Renilla reniformis; differential gene expression; protein localisation;
 KW gene expression tracking; fluorescence.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 66 /note= "Y66H"
 FT Misc-difference 145 /note= "Y145F"
 FT
 XX
 FN W09623810-A1.
 XX
 PD 08-AUG-1996.
 XX
 PF 13-NOV-1995; 95WO-US014692.
 XX
 PR 10-NOV-1994; 94US-00337915.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Tsien RY, Heim R;
 XX
 DR WPI; 1996-371370/37.
 XX
 PT New modified Aequorea green fluorescent polypeptide(s) - having amino
 acid changes to provide prods. which exhibit different excitation and
 emission spectra.
 XX
 PS Claim 10; Page ?; 39pp; English.
 XX
 CC AAW05305-W05312 represent Aequorea victoria (jellyfish) green fluorescent
 protein (GFP) of the invention (see AAW05304 for the wild type protein).
 CC The fluorescence of this protein is generated by cyclisation and
 CC oxidation of the Ser-Tyr-Gly sequence at positions 65-67. Aequorea GFP
 CC has two absorbance peaks, as opposed to the one absorbance peak seen in
 CC the related GFP from the sea pansy (Renilla reniformis). The
 CC modifications present in these sequences were created by subjecting the
 CC cDNA encoding this sequence to site directed mutagenesis using mutagenic
 CC PCR primers, or hydroxylamine treatment. These GFPs of the invention are
 CC modified to lead to the formation of products with markedly different
 CC excitation and emission spectra. Visibly distinct colours, and increased
 CC intensities of emission make these products useful in a wide variety of
 CC contexts, such as tracking of differential gene expression and protein
 CC localisation. The mutations can also be created to modify the encoded GFP
 CC so that it only possesses one absorbance peak
 XX
 SQ Sequence 238 AA;
 Query Match 98.1%; Score 1241; DB 2; Length 238;
 Best Local Similarity 97.9%; Pred. No. 1.2e-120;
 Matches 233; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPIVLVDGVDGNGHFKFSVSGEGDATYGKLTILKFICTTGKLPVPPPTL 60
 Db 1 MSKGEELFTGVVPIVLVDGVDGNGHFKFSVSGEGDATYGKLTILKFICTTGKLPVPPPTL 60
 QY 61 VTFSSGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLV 120
 Db 61 VTFSSGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNKYKTRAEVKEGDTLV 120
 QY 121 NRLEKIGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 Db 121 NRLEKIGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 QY 181 HYQONTPIGDGVPVLLPDNHYLSTQSALS KDPKPKRDHVMVLEFVTAAGITHGMDELYK 238
 Db 181 HYQONTPIGDGVPVLLPDNHYLSTQSALS KDPKPKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 9
 ABG76005
 ID ABG76005 standard; protein; 238 AA.
 XX
 AC ABG76005;
 XX
 DT 30-APR-2003 (first entry)
 XX
 DE Jellyfish GFP mutant P4-3.
 XX
 KW Jellyfish; enzyme; green fluorescent protein; GFP; FRET; mutant;
 KW fluorescent resonance energy transfer; tandem fluorescent protein;
 KW enzymatic assay; Alzheimer's disease; hypertension; inflammation;
 KW apoptosis; AIDS; acquired immunodeficiency syndrome; mutin.
 XX
 OS Aequorea victoria.
 OS Synthetic.
 XX
 XX
 Key Location/Qualifiers
 FT Misc-difference 66 /note= "Wild-type Tyr substituted by His"
 FT Misc-difference 145 /note= "Wild-type Tyr substituted by Phe"
 FT
 XX
 XX US2002164674-A1.
 XX
 XX 07-NOV-2002.
 XX
 XX 25-JAN-2002; 2002US-00057505.
 XX
 XX 31-JAN-1996; 96US-00594575.
 PR 31-JAN-1997; 97US-00792553.
 PR 13-SEP-1999; 99US-00396003.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX
 PI Tsien RY, Heim R, Cubitt A;
 XX
 XX WPI; 2003-247255/24.
 XX
 XX New tandem fluorescent protein construct comprising a donor or acceptor
 FT fluorescent protein moiety or a linker moiety that couples the donor and
 FT acceptor moieties, useful in enzymatic assays.
 XX
 XX Claim 3; Page; 34pp; English.
 XX
 XX The invention relates to a tandem fluorescent protein construct
 CC comprising a donor or acceptor fluorescent protein moiety (e.g. jellyfish
 CC green fluorescent protein, GFP) or a linker moiety that couples the donor
 CC and acceptor moieties. Also include are a recombinant nucleic acid coding
 CC for expression of the tandem fluorescent protein construct, an expression
 CC vector comprising expression control sequences operatively linked to a
 CC sequence coding for the expression of the tandem fluorescent protein
 CC construct, a host cell transfected with the expression vector,
 CC determining whether a sample contains an enzyme or whether a compound
 CC alters the activity of an enzyme, determining the amount of activity of
 CC an enzyme in a cell and testing for cleavage enzyme activity. The tandem
 CC fluorescent protein construct is useful in enzymatic assays, using the
 CC principle of fluorescent resonance energy transfer (FRET) between the
 CC donor and acceptor moieties. The tandem fluorescent proteins are
 CC particularly useful in assays for protease activity. Proteases play an
 CC essential role in many disease processes e.g. Alzheimer's disease,
 CC hypertension, inflammation, apoptosis and AIDS (acquired immunodeficiency
 CC syndrome). Tandem fluorescent proteins were constructed comprising
 CC mutants of GFP with altered fluorescent spectra. The present sequence
 CC represents a mutant GFP suitable for incorporation into a tandem
 CC fluorescent protein of the invention. Note: The present sequence is not
 CC shown in the specification but was created by the indexer using the
 CC information in the specification and the wild-type GFP protein appearing
 CC as ABG75980
 XX
 XX Sequence 238 AA;

Query Match 98.1%; Score 1241; DB 6; Length 238;
 Best Local Similarity 97.9%; Pred. No. 1.2e-120;
 Matches 233; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGDGYGKLTFLKFICTTGKLPVPWPTL 60
 DB 1 MSKGEELFTGVVPILVELDGDVNGHKFSVSGEGDGYGKLTFLKFICTTGKLPVPWPTL 60
 QY 61 VTFESGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
 DB 61 VTFESGVQCFSRYPDHMKRHDFFKSAPEGVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
 QY 121 NRILKGIIDFKEDGNILGHKLEYFINSYVITADKQNGIKANFKIRHNTEDGSVOLAD 180
 DB 121 NRILKGIIDFKEDGNILGHKLEYFINSYVITADKQNGIKANFKIRHNTEDGSVOLAD 180
 QY 181 HYQONTPIGDGPFVLLPDNHYLSTQALSQKDPKRDHMLLEFVTAAGITHGMDELYK 238
 DB 181 HYQONTPIGDGPFVLLPDNHYLSTQALSQKDPKRDHMLLEFVTAAGITHGMDELYK 238
 RESULT 10
 AAE34991
 ID AAE34991 standard; protein; 238 AA.
 XX
 AC AAE34991;
 XX
 DT 28-MAY-2003 (first entry)
 XX
 DE Aequorea victoria mutant green fluorescent protein (N146I).
 XX
 KW Phosphorylation indicator; fluorescent protein; detection; phosphatase;
 KW kinase; green fluorescent protein; GFP; mutant; mutin.
 XX
 OS Aequorea victoria.
 OS Synthetic.
 XX
 XX
 Key Location/Qualifiers
 FT Misc-difference 146 /note= "Wild-type Asn substituted with Ile"
 FT
 XX
 XX WO200295058-A2.
 XX
 XX 28-NOV-2002.
 XX
 XX 24-MAY-2002; 2002WO-US016955.
 XX
 XX 24-MAY-2001; 2001US-00865291.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX
 XX Tsien RY, Ting AY, Zhang J;
 XX WPI; 2003-148474/14.
 XX
 XX Novel chimeric phosphorylation indicators, useful for detecting
 FT kinase/phosphatase in samples, has donor molecule, phosphorylatable
 FT domain, phosphoaminoacid binding domain, and acceptor molecule, in
 FT operative linkage.
 XX
 XX Disclosure; Col; 38pp; English.
 XX
 XX The present invention relates to chimeric phosphorylation indicators
 CC comprising a phosphorylation polypeptide and a fluorescent protein or
 CC operative linkage, a donor molecule, a phosphorylatable domain, a
 CC phosphoaminoacid binding domain (PABD), and an acceptor molecule. The
 CC phosphorylation indicators of the invention are useful for detecting in
 CC kinases or phosphatases in a biological sample. They are also useful in
 CC high throughput analysis e.g. for detecting a kinase inhibitor or
 CC phosphatase inhibitor. The present sequence is Aequorea victoria mutant
 CC green fluorescent protein (GFP; N146I) used in the invention. Note: This
 CC sequence is not shown in the specification but is derived from Aequorea
 CC victoria wild-type GFP shown as SEQ ID NO: 2 in column 53-54 of the

```
CC specification (AAE34957)
XX Sequence 238 AA;
SQ

Query Match      98.1%; Score 1241; DB 6; Length 238;
Best Local Similarity 97.5%; Pred. No. 1.2e-120;
Matches 232; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLVDGVDVNGHKFSVSGEGEDATYGKLTLCFTCTTGKLPVPWPTL 60
Db 1 MSKGELFTGVVPIVLVDGVDVNGHKFSVSGEGEDATYGKLTLCFTCTTGKLPVPWPTL 60
QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGVQVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKQHDFFKSAMPEGVQVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNHYITADKQNGIKANFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNHYITADKQNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKPKRDHVMVLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKPKRDHVMVLEFVTAAGITHGMDELYK 238

RESULT 11
AAW31880
ID AAW31880 standard; protein; 501 AA.
XX
AC AAW31880;
XX
DT 03-FEB-1998 (first entry)
XX
DE GFP variants P4-3 and W7 tandem fluorescent protein construct.
XX
KW North West Pacific jellyfish; green fluorescent protein; GFP; P4-3;
KW fluorescent resonance energy transfer; FRET; enzymatic assay; W7;
KW enzyme inhibitor; enzyme promoter; PCR primer; protease cleavage site;
KW Tandem fluorescent protein construct; blue fluorescent protein.
XX
OS Synthetic.
OS Aequorea victoria.
XX
FH Key Location/Qualifiers
FT Protein 1..238
FT /label= P4-3_GFP_variant
FT Misc-difference 66
FT /label= Y66H
FT /note= "wild type Tyr substituted with His"
FT Misc-difference 145
FT /label= Y145F
FT /note= "wild type Tyr substituted with Phe"
FT Peptide 239..264
FT /label= linker_moeity
FT Cleavage-site 250..251
FT /label= trypsin_cleavage_site
FT Cleavage-site 253..254
FT /label= calpain_cleavage_site
FT Cleavage-site 258..259
FT /label= trypsin_enterokinase_cleavage_site
FT Protein 265..501
FT Misc-difference 329
FT /label= W7_GFP_variant
FT /label= Y66W
FT /note= "wild type Tyr substituted with Trp"
FT Misc-difference 409
FT /label= N146L
FT /note= "wild type Asn substituted with Leu"
FT Misc-difference 416
FT /label= M153T
FT /note= "wild type Met substituted with Thr"
FT Misc-difference 426
FT /label= V163A
```

```
FT Misc-difference 475
FT /label= N212K
FT /note= "wild type Val substituted with Ala"
XX
PN WO9728261-A1.
XX
PD 07-AUG-1997.
XX
PF 31-JAN-1997; 97WO-US001457.
XX
PR 31-JAN-1996; 96US-00594575.
XX
PA (REGC ) UNIV CALIFORNIA.
PA (AURO-) AURORA BIOSCIENCES CORP.
XX
PI Tsien RY, Heim R, Cubitt A;
XX
DR MPI; 1997-402615/37.
XX
Tandem fluorescent protein constructs - have donor and acceptor moieties
exhibiting fluorescent linked via cleavable peptide linker, useful in
enzymatic assays.
XX
PS Example 1; Page; 89pp; English.
XX
CC This protein sequence is that of a novel tandem fluorescent protein
construct, made using Aequorea victoria (North West Pacific jellyfish)
green fluorescent protein (GFP) variants P4-3 and W7. Both of these
fluoresce at a shorter wavelength than GFP. The construct comprises a
donor (e.g. P4-3) and an acceptor (e.g. W7) fluorescent protein moiety
(donor and acceptors can be green or blue fluorescent proteins), and a
linker coupling them. Preferably, the donor is positioned at the N-
terminus of the polypeptide relative to the acceptor. The linker moiety
is a peptide 5-50 amino acids in length containing a protease cleavage
site. In this example, the linker moiety contains many recognition sites
for proteases, including trypsin, calpain and enterokinase. The donor and
acceptor moieties exhibit fluorescent resonance energy transfer (FRET)
when they are cleaved. The constructs are used in enzymatic assays and
can be used to isolate new enzymes or enzyme inhibitors or promoters. The
specific activity of enzyme (in vivo and in vitro) and compounds altering
enzyme activity can be obtained. FRET and hence activity of specific
compounds is measured from the acceptor or donor moiety or maybe obtained
using a ratio between the two
XX
SQ Sequence 501 AA;
```

```
Query Match      98.1%; Score 1241; DB 2; Length 501;
Best Local Similarity 97.9%; Pred. No. 3.7e-120;
Matches 233; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGELFTGVVPIVLVDGVDVNGHKFSVSGEGEDATYGKLTLCFTCTTGKLPVPWPTL 60
Db 1 MSKGELFTGVVPIVLVDGVDVNGHKFSVSGEGEDATYGKLTLCFTCTTGKLPVPWPTL 60
QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGVQVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSHGVQCFSRYPDHMKRHDFFKSAMPEGVQVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFIHNHYITADKQNGIKANFKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFIHNHYITADKQNGIKANFKIRHNIEDGSVQLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKPKRDHVMVLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKPKRDHVMVLEFVTAAGITHGMDELYK 238
```

```
RESULT 12
AAW31876
ID AAW31876 standard; protein; 514 AA.
XX
AC AAW31876;
```

XX 03-FEB-1998 (first entry)

DE GFP variants S65C and P4-3 tandem fluorescent protein construct.

XX North West Pacific jellyfish; green fluorescent protein; GFP; S65C;
 KW fluorescent resonance energy transfer; FRET; enzymatic assay; P4-3;
 KW enzyme inhibitor; enzyme promoter; PCR primer; protease cleavage site;
 XX Tandem fluorescent protein construct; blue fluorescent protein.

XX Synthetic.
 OS Aequorea victoria.

XX Key Location/Qualifiers

FT Peptide 5..10
 FT /label= polyhistidine_tag
 FT Protein 14..251
 FT /label= S65C_GFP_variant
 FT Misc-difference 78
 FT /label= S65C
 FT /note= "wild type Ser substituted with Cys"
 FT Peptide 252..276
 FT /label= linker_moiety
 FT Cleavage-site 262..263
 FT /label= trypsin_cleavage_site
 FT Cleavage-site 265..266
 FT /label= calpain_cleavage_site
 FT Cleavage-site 270..271
 FT /label= trypsin_enterokinase_cleavage_site
 FT Protein 277..514
 FT /label= P4-3_GFP_variant
 FT Misc-difference 342
 FT /label= Y66H
 FT /note= "wild type Tyr substituted with His"
 FT Misc-difference 421
 FT /label= Y145F
 FT /note= "wild type Tyr substituted with Phe"

XX WO9728261-A1.

XX 07-AUG-1997.

XX 31-JAN-1997; 97WO-US001457.

XX 31-JAN-1996; 96US-00594575.

XX (REGC) UNIV CALIFORNIA.
 XX (AURO-) AURORA BIOSCIENCES CORP.

XX Tsien RY, Heim R, Cubitt A;
 WPI; 1997-402615/37.

XX Tandem fluorescent protein constructs - have donor and acceptor moieties
 PT exhibiting fluorescent linked via cleavable peptide linker, useful in
 FT enzymatic assays.

XX Claim 3; Page; 88pp; English.

XX This protein sequence is that of a novel tandem fluorescent protein
 CC construct, made using Aequorea victoria (North West Pacific jellyfish)
 CC green fluorescent protein (GFP) variants S65C and P4-3. P4-3 fluoresces
 CC at a shorter wavelength than GFP. The construct comprises a donor (e.g.
 CC S65C) and an acceptor (e.g. P4-3) fluorescent protein moiety (donors and
 CC acceptors can be green or blue fluorescent proteins), and a linker
 CC coupling them. Preferably, the donor is positioned at the N-terminus of
 CC the polypeptide relative to the acceptor. The linker moiety is a peptide
 CC 5-50 amino acids in length containing a protease cleavage site. In this
 CC example, the linker moiety contains many recognition sites for proteases,
 CC including trypsin, calpain and enterokinase. The donor and acceptor
 CC moieties exhibit fluorescent resonance energy transfer (FRET) when they
 CC are cleaved. The constructs are used in enzymatic assays and can be used
 CC to isolate new enzymes or enzyme inhibitors or promoters. The specific

CC activity of enzyme (in vivo and in vitro) and compounds altering enzyme
 CC activity can be obtained. FRET and hence activity of specific compounds
 CC is measured from the acceptor or donor moiety or maybe obtained using a
 CC ratio between the two. Note: The present sequence does not appear in the
 CC specification, it has been made by modifying the native GFP sequence, and
 CC adding the linker moiety in the appropriate place

XX Sequence 514 AA;
 SQ

Query Match 98.1%; Score 1241; DB 2; Length 514;
 Best Local Similarity 97.9%; Pred. No. 3.8e-120;
 Matches 233; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKGLTLKFICTTGKLPVPWPTL 60
 Db 277 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYKGLTLKFICTTGKLPVPWPTL 336

OY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
 Db 337 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 396

OY 121 NRIELKGIDFKEDGNILGHKLEYNFIHSNVTITADKOKNGIKANFKIRHNIEDGSVQLAD 180
 Db 397 NRIELKGIDFKEDGNILGHKLEYNFIHSNVTITADKOKNGIKANFKIRHNIEDGSVQLAD 456

OY 181 HYQONTPIGDGPFVLLPDNHYLSTQSALSKDPKPKRDHMLLEFVTAAGITHGMDELYK 238
 Db 457 HYQONTPIGDGPFVLLPDNHYLSTQSALSKDPKPKRDHMLLEFVTAAGITHGMDELYK 514

RESULT 13
 AAB73553
 ID AAB73553 standard; protein; 238 AA.
 XX AC AAB73553;
 XX DT 07-AUG-2001 (first entry)
 XX DE Green fluorescent protein variant GFPuv.
 XX KW Phenotype selection; non-selectable; fusion protein; stable expression;
 KW selectable marker; antibiotic resistance gene; Escherichia coli;
 KW green fluorescent protein; GFP; GFPuv; mutant; mutein.
 XX OS Aequorea victoria.
 XX OS Synthetic.
 XX PN WO200129225-A1.
 XX PD 26-APR-2001.
 XX PF 29-MAR-2000; 2000WO-US008477.
 XX PR 21-OCT-1999; 99US-0160461P.
 XX PR 22-FEB-2000; 2000US-00510097.
 XX PA (PANO-) PANORAMA RES INC.
 XX PI Balint RF;
 XX DR WPI; 2001-282162/29.
 XX DR N-PSDB; AAH20246.

PT Obtaining cells expressing mutant protein, comprises selecting from cells
 PT transformed with library of mutagenized protein coding sequences joined
 PT to selector protein, which confers growth under selective conditions.
 XX Example 2; Page; 52pp; English.

XX The invention relates to methods whereby variants of proteins which do
 CC not confer selectable phenotypes can be selected for stable expression in
 CC host cells (especially Escherichia coli). The methods can be used to
 CC obtain mutants of a desired protein optimised for expression in the host

CC cells, for obtaining a mutant protein with enhanced stability relative to
 CC the wild-type protein, and for identifying peptides that stabilise an
 CC unstable protein. The methods all involve expressing the protein of
 CC interest as a fusion with a protein that can confer a selectable
 CC phenotype, such as an antibiotic resistance protein. The transformed host
 CC cells are then grown under selective pressure (e.g., presence of
 CC antibiotic). The cells able to grow under such conditions are those which
 CC contain fusion proteins which are optimised for expression or which are
 CC more stable, as this property will also correlate with an increased
 CC amount or rate of synthesis of the selectable marker. The invention also
 CC discloses mutants of green fluorescent protein (GFP) selected for
 CC increased stability using the method of the invention. The present
 CC sequence represents a GFP variant (GFPuv) as encoded by plasmid pGFPuv
 CC (Genbank U62636), which was used in an exemplification of the invention.
 CC The present sequence is not shown in the specification, but was derived
 CC from Genbank accession number U62636

XX SQ Sequence 238 AA;

Query Match 98.0%; Score 1240; DB 4; Length 238;
 Best Local Similarity 97.9%; Pred. No. 1.6e-120;
 Matches 233; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
 Db 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
 QY 61 VTTFSXGVCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
 Db 61 VTTFSXGVCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKRDHMLLEFVTAAGITHGMDELYK 238
 Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 14

AAE34992
 ID AAE34992 standard; protein; 238 AA.

XX AAE34992;

AC AC
 DT 28-MAY-2003 (first entry)

XX Aequorea victoria mutant green fluorescent protein (M153T).

XX Phosphorylation indicator; fluorescent protein; detection; phosphatase;
 XX kinase; green fluorescent protein; GFP; mutant; mutin.

XX Aequorea victoria.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 153 /note= "Wild-type Met substituted with Thr"

XX WO200295058-A2.

XX 28-NOV-2002.

XX 24-MAY-2002; 2002WO-US016955.

XX 24-MAY-2001; 2001US-00865291.

XX (REGC) UNIV CALIFORNIA.

XX Tsien RY, Ting AY, Zhang J;

XX WPI; 2003-148474/14.

XX Novel chimeric phosphorylation indicators, useful for detecting
 PT kinase/phosphatase in samples, has donor molecule, phosphorylatable
 PT domain, phosphoaminoacid binding domain, and acceptor molecule, in
 PT operative linkage.

XX Disclosure; Col; 38pp; English.

XX The present invention relates to chimeric phosphorylation indicators
 CC comprising a phosphorylation polypeptide and a fluorescent protein or in
 CC operative linkage, a donor molecule, a phosphorylatable domain, a
 CC phosphoaminoacid binding domain (PABD) and an acceptor molecule. The
 CC phosphorylation indicators of the invention are useful for detecting
 CC kinases or phosphatases in a biological sample. They are also useful in
 CC high throughput analysis e.g. for detecting a kinase inhibitor or
 CC phosphatase inhibitor. The present sequence is Aequorea victoria mutant
 CC green fluorescent protein (GFP; M153T) used in the invention. Note: This
 CC sequence is not shown in the specification but is derived from Aequorea
 CC victoria wild-type GFP shown as SEQ ID NO: 2 in column 53-54 of the
 CC specification (AAE34992/)

XX SQ Sequence 238 AA;

Query Match 98.0%; Score 1240; DB 6; Length 238;
 Best Local Similarity 97.5%; Pred. No. 1.6e-120;
 Matches 232; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
 Db 1 MSKGEELFTGVVPILVELDGVNKGKFSVSGEGEDATYGLTKLFICTTGKLPVWPPTL 60
 QY 61 VTTFSXGVCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
 Db 61 VTTFSXGVCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 Db 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVIITADKQKNGIKANFKIRHNIEDGSVOLAD 180
 QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKRDHMLLEFVTAAGITHGMDELYK 238
 Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALSKDPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 15

AAW05311
 ID AAW05311 standard; protein; 238 AA.

XX AAW05311;

XX 02-APR-1997 (first entry)

XX GFP mutant Y66W/I123V/Y145H/H148R/M153T/V163A/N212K.

XX Green fluorescent protein; GFP; jellyfish; Aequorea victoria; sea pansy;
 KW Renilla reniformis; differential gene expression; protein localisation;
 KW gene expression tracking; fluorescence.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 66 /note= "Y66W"

FT Misc-difference 123 /note= "I123V"

FT Misc-difference 145 /note= "Y145V"

FT Misc-difference 148 /note= "H148R"

FT Misc-difference 153 /note= "M153T"

FT Misc-difference 163 /note= "V163A"

FT Misc-difference 212
FT /note= "N212K"
XX
XX
XX WO9623810-A1.
XX
XX 08-AUG-1996.
XX
XX 13-NOV-1995; 95WO-US014692.
XX
XX 10-NOV-1994; 94US-00337915.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Tsien RY, Heim R;
XX
XX WPI; 1996-371370/37.
XX
XX New modified Aequorea green fluorescent polypeptide(s) - having amino
XX acid changes to provide prods. which exhibit different excitation and
XX emission spectra.
XX
XX Claim 12; Page ?; 39pp; English.
XX
XX AAW05305-W05312 represent Aequorea victoria (jellyfish) green fluorescent
XX protein (GFP) of the invention (see AAW05304 for the wild type protein).
XX The fluorescence of this protein is generated by cyclisation and
XX oxidation of the Ser-Tyr-Gly sequence at positions 65-67. Aequorea GFP
XX has two absorbance peaks, as opposed to the one absorbance peak seen in
XX the related GFP from the sea pansy (Renilla reniformis). The
XX modifications present in these sequences were created by subjecting the
XX cDNA encoding this sequence to site directed mutagenesis using mutagenic
XX PCR primers, or hydroxylamine treatment. These GFPs of the invention are
XX modified to lead to the formation of products with markedly different
XX excitation and emission spectra. Visibly distinct colours, and increased
XX intensities of emission make these products useful in a wide variety of
XX contexts, such as tracking of differential gene expression and protein
XX localisation. The mutations can also be created to modify the encoded GFP
XX so that it only possesses one absorbance peak
XX
SQ Sequence 238 AA;

Query Match 97.9%; Score 1239; DB 2; Length 238;
Best Local Similarity 97.9%; Pred. No. 2e-120;
Matches 233; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGDATYKGLTKLKFICTTCKLPVWPPTL 60
Db 1 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGDATYKGLTKLKFICTTCKLPVWPPTL 60

Qy 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNYKTRAEVKPEGDTLV 120
Db 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAPEGYVQERTIFFKDDGNYKTRAEVKPEGDTLV 120

Qy 121 NRLEKGDIDKEDGNILGHKLEYNFIHNVIYITADKQNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRLEKGDIDKEDGNILGHKLEYNFIHNVIYITADKQNGIKANFKIRHNIEDGSVOLAD 180

Qy 181 HYQONTPIGGPVLPPNNHYLSTQSALS KDPKRDHMLLEFVTAAGITGMDELYK 238
Db 181 HYQONTPIGGPVLPPNNHYLSTQSALS KDPKRDHMLLEFVTAAGITGMDELYK 238

Search completed: April 19, 2004, 15:05:58
Job time : 61 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 14:56:35 ; Search time 17 Seconds
(without alignments)
728.982 Million cell updates/sec

Title: US-10-057-505-2-COPY

Perfect score: 1265

Sequence: 1 MSKGEELTGVPVILVELDGL.....VLLEFVTAAGITHGMDLYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1234	97.5	238	1	GFP_AEQVI
2	87.5	6.9	357	1	TRMA_CAMJE
3	87.5	6.9	613	1	PEPF_MYCPU
4	85.5	6.8	2222	1	DPOE_YEAST
5	85	6.7	689	1	AC2L_HUMAN
6	84.5	6.7	501	1	AMPA_WIGER
7	84.5	6.7	886	1	ITH3_MESAU
8	82.5	6.5	879	1	SYL_XYLFA
9	82	6.5	795	1	D152_HAEIN
10	82	6.5	797	1	D151_HAEIN
11	81	6.4	366	1	SET7_HUMAN
12	81	6.4	649	1	ACS2_RHIME
13	81	6.4	1312	1	STRH_STREN
14	80.5	6.4	533	1	CP51_CANGA
15	80.5	6.4	533	1	NIFD_CLOPA
16	80.5	6.4	700	1	CAN2_HUMAN
17	80.5	6.4	700	1	CAN2_MACPA
18	80	6.3	439	1	SV62_DILSON
19	80	6.3	462	1	K6EF_METOJA
20	80	6.3	504	1	YC03_KLEPN
21	80	6.3	793	1	D153_HAEIN
22	80	6.3	874	1	SLAP_BACLI
23	79.5	6.3	274	1	SPED_CLOAB
24	79.5	6.3	422	1	CAN2_RABIT
25	79.5	6.3	879	1	SYL_XYLFT
26	79	6.2	400	1	TRMB_HELPJ
27	79	6.2	682	1	AC2L_MOUSE
28	79	6.2	1164	1	BAG_STRAG
29	79	6.2	1802	1	HK31_YEAST
30	78.5	6.2	530	1	TP63_SULSH
31	78.5	6.2	887	1	ITH3_RAT
32	78.5	6.2	1127	1	LONH_PYRHO
33	78.5	6.2	1224	1	CYL_HUMAN

RESULT 1
GFP_AEQVI
ID GFP_AEQVI STANDARD; PRT; 238 AA.
AC P42212; Q17104; Q27903;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92175527; PubMed=1347277;
RA Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G.,
RA Cormier M.J.;
RT "Primary structure of the Aequorea victoria green-fluorescent
RT protein.";
RL Gene 111:229-233 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94185810; PubMed=8137953;
RA Inouye S., Tsuji F.I.;
RT "Aequorea green fluorescent protein. Expression of the gene and
RT fluorescence characteristics of the recombinant protein.";
RL FEBS Lett. 341:277-280 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97299832; PubMed=9154981;
RA Rouwendal G.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;
RT "Enhanced expression in tobacco of the gene encoding green fluorescent
RT protein by modification of its codon usage.";
RL Plant Mol. Biol. 33:989-999 (1997).
RN [4]
RP CHROMOPHORE.
RX MEDLINE=93192221; PubMed=8448132;
RA Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;
RT "Chemical structure of the hexapeptide chromophore of the Aequorea
RT green-fluorescent protein.";
RL Biochemistry 32:1212-1218 (1993).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96355665; PubMed=8703075;
RA Ormoe M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,
RA Remington S.J.;
RT "Crystal structure of the Aequorea victoria green fluorescent
RT protein.";
RL Science 273:1392-1395 (1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=98294543; PubMed=9631087;
RA Yang F., Moss L.G., Phillips G.N. Jr.;
RT "The molecular structure of green fluorescent protein.";
RL Nat. Biotechnol. 14:1246-1251 (1996).

34 78 6.2 207 1 CAN2_BOVIN
35 78 6.2 353 1 HIS7_BUCAL
36 78 6.2 488 1 MSRA_STAPB
37 78 6.2 806 1 SYL_UREPA
38 77.5 6.1 324 1 CAN2_PIG
39 77.5 6.1 355 1 PLK_CHICK
40 77.5 6.1 1224 1 COFA_BOVIN
41 77.5 6.1 758 1 PEPX_STRMU
42 77 6.1 414 1 RCA_ANASP
43 77 6.1 652 1 ACSA_AGRRH
44 77 6.1 817 1 RPOD_ASTLO
45 77 6.1 858 1 KCB1_HUMAN
Q27971 bos taurus
P57203 buchnera ap
P23212 staphylococ
Q99GC0 usapiasma
P43367 sus scrofa
P07354 gallus gall
Q84VS2 streptococ
Q27954 bos taurus
P58555 anabaena sp
Q9KWA3 agrobacteri
P58132 astasia lon
Q14721 homo sapien

ALIGNMENTS

[7]
 RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMISSION.
 RX MEDLINE=9845509; PubMed=9782051;
 RA Wachter R.M., Elsiger M.A., Kallio K., Hanson G.T., Remington S.J.;
 RT "Structural basis of spectral shifts in the yellow-emission variants
 RL of green fluorescent protein.";
 RL Structure 6:1267-1277(1998).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=99238303; PubMed=10220315;
 RA Elsiger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
 RT "Structural and spectral response of green fluorescent protein
 RL variants to changes in pH.";
 RL Biochemistry 38:5296-5301(1999).
 CC -!- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
 CC blue chemiluminescence of the protein aequorin into green
 CC fluorescent light by energy transfer. Fluoresces in vivo upon
 CC receiving energy from the Ca(2+)-activated photoprotein aequorin.
 CC Absorbs light maximally at 395 nm and exhibits a smaller
 CC absorbance peak at 470 nm. The fluorescence emission spectrum
 CC peaks at 509 nm with a shoulder at 540 nm.
 CC -!- SUBUNIT: Monomer.
 CC -!- TISSUE SPECIFICITY: Photocytes.
 CC -!- PTM: Contains a covalently attached chromophore, which is composed
 CC of modified amino acid residues. The chromophore is formed upon
 CC cyclization of the residues Ser-dehydrotyr-Gly.
 CC -!- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
 CC chimeric proteins of GFP linked to other proteins where it
 CC functions as a fluorescent protein tag. GFP tolerates N- and C-
 CC terminal fusion to a broad variety of proteins. It has been
 CC expressed in bacteria, yeast, slime mold, plants, Drosophila,
 CC zebrafish, and in mammalian cells. As a noninvasive fluorescent
 CC marker in living cells, it allows for a wide range of applications
 CC where it may function as a cell lineage tracer, reporter of gene
 CC expression, or as a measure of protein-protein interactions.
 CC -!- DATABASE: NAME-Protein Spotlight;
 CC NOTE=Issue 11 of June 2001;
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt011.html".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M62654; AAA27722.1; -;
 CC EMBL; M62653; AAA27721.1; -;
 CC EMBL; L29345; AAA58246.1; -;
 CC EMBL; X96418; CAA65278.1; -;
 CC FIR; JS0692; JQ1514.
 CC PDB; 1B9C; 17-NOV-00.
 CC PDB; 1BFF; 07-JUL-97.
 CC PDB; 1C4F; 14-JUN-00.
 CC PDB; 1EWA; 08-NOV-96.
 CC PDB; 1EMB; 16-JUN-97.
 CC PDB; 1EMC; 20-AUG-97.
 CC PDB; 1EME; 20-AUG-97.
 CC PDB; 1EMF; 20-AUG-97.
 CC PDB; 1EMG; 12-MAY-99.
 CC PDB; 1EMK; 20-AUG-97.
 CC PDB; 1EMI; 20-AUG-97.
 CC PDB; 1EMW; 20-AUG-97.
 CC PDB; 1F09; 17-NOV-00.
 CC PDB; 1F0B; 17-NOV-00.
 CC PDB; 1F0J; 17-NOV-00.
 CC PDB; 1HGT; 11-JAN-97.
 CC PDB; 1HGI; 15-JAN-02.
 CC PDB; 1HUY; 04-JUL-01.
 CC PDB; 1JBY; 07-JAN-03.
 CC PDB; 1JBZ; 07-JAN-03.
 CC PDB; 1KPS; 28-AUG-02.
 CC PDB; 1KYP; 10-APR-02.

PDB; 1KYR; 10-APR-02.
 PDB; 1KYS; 10-APR-02.
 PDB; 1VFP; 28-OCT-98.
 PDB; 2EMD; 20-AUG-97.
 PDB; 2EMN; 20-AUG-97.
 PDB; 2EMO; 20-AUG-97.
 DR InterPro; IPR009017; GFP-like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFPUNSCENT.
 DR ProDom; PR013756; Green_fl_protein; 1.
 KW Luminescence; 3D-structure.
 FT CROSSLINK 65 67 5-imidazolinone (Ser-Gly).
 FT MOD RES 66 66 2,3-DIHYDROTYROSINE.
 FT VARIANT 100 100 F -> Y.
 FT VARIANT 108 108 T -> S.
 FT VARIANT 141 141 L -> M.
 FT VARIANT 219 219 V -> I.
 FT CONFLICT 2 2 S -> G (IN REF. 3).
 FT CONFLICT 25 25 H -> Q (IN REF. 2).
 FT CONFLICT 80 80 Q -> R (IN REF. 3).
 FT CONFLICT 157 157 Q -> P (IN REF. 2).
 FT CONFLICT 172 172 E -> K (IN REF. 2).
 FT HELIX 4 8
 FT STRAND 12 22
 FT TURN 23 24
 FT STRAND 25 36
 FT TURN 37 40
 FT STRAND 41 48
 FT TURN 49 50
 FT HELIX 57 60
 FT TURN 61 63
 FT HELIX 69 71
 FT STRAND 73 73
 FT HELIX 76 81
 FT HELIX 83 86
 FT TURN 87 90
 FT STRAND 92 100
 FT TURN 101 102
 FT STRAND 105 115
 FT TURN 116 117
 FT STRAND 118 128
 FT TURN 132 133
 FT TURN 135 139
 FT STRAND 141 141
 FT STRAND 148 155
 FT TURN 156 159
 FT STRAND 160 171
 FT TURN 172 173
 FT STRAND 176 187
 FT STRAND 199 208
 FT TURN 211 212
 FT STRAND 217 227
 SQ SEQUENCE 238 AA; 26886 MW; EA5A6F21FBFB6E05 CRC64;
 Query Match 97.5%; Score 1234; DB 1; Length 238;
 Best Local Similarity 97.1%; Pred. No. 2.3e-95;
 Matches 231; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 MSKGELEFTGVVPIVLVDGVNGHFKFSVSGEGDGYGKLTGKLCFCTTGKLPVWPPTL 60
 DB 1 MSKGELEFTGVVPIVLVDGVNGHFKFSVSGEGDGYGKLTGKLCFCTTGKLPVWPPTL 60
 QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKXAMPEGVGVQERTIFFKDDGNYKTRAEVYKFGDTLV 120
 DB 61 VTTFSYGVQCFSRYPDHMKQHDFFKXAMPEGVGVQERTIFFKDDGNYKTRAEVYKFGDTLV 120
 QY 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSGVOLAD 180
 DB 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQNGIKANFKIRHNIEDGSGVOLAD 180
 QY 181 HYQONTPTIGDGPVLLPDNHYLSLTQALSQDKPKERDHWLVLEFVTAAGTTHGMDELYK 238

[illegible]

01-AUG-1991 (Rel. 19, Created)
 01-AUG-1991 (Rel. 19, Last sequence update)
 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA polymerase epsilon, catalytic subunit A (EC 2.7.7.7) (DNA polymerase II subunit A)
 DE POL2 OR DUN2 OR YNL262W OR N0825.
 GN Saccharomyces cerevisiae (Baker's yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1214-1221.
 RC MEDLINE=90381771; PubMed=2169349;
 RA Morrison A., Araki H., Clark A.B., Hamatake R.K., Sugino A.;
 RT "A third essential DNA polymerase in *S. cerevisiae*.";
 RL Cell 62:1143-1151(1990).
 RN [2]
 RP SEQUENCE OF 1-2221 FROM N.A.
 RC STRAIN=S288c / FY1679;
 RA MEDLINE=96310631; PubMed=8740425;
 RA Sen-Gupta M., Lyck R., Fleig U., Niedenthal R.K., Hegemann J.H.;
 RT "The sequence of a 24,152 bp segment from the left arm of chromosome XIV from *Saccharomyces cerevisiae* between the BNI1 and the POL2 genes.";
 RL Yeast 12:505-514(1996).
 RN [3]
 RP TEMPERATURE SENSITIVE MUTANTS
 RC MEDLINE=92164663; PubMed=1537345;
 RA Araki H., Ropp P.A., Johnson A.L., Johnston L.H., Morrison A., Sugino A.;
 RT "DNA polymerase II, the probable homolog of mammalian DNA polymerase epsilon, replicates chromosomal DNA in the yeast *Saccharomyces cerevisiae*.";
 RL EMBO J. 11:733-740(1992).
 RN
 CC -!- FUNCTION: DNA POLYMERASE II PARTICIPATES IN CHROMOSOMAL DNA REPLICATION.
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate + {DNA}(N).
 CC -!- SUBUNIT: CONSISTS OF FIVE SUBUNITS (200 kDa, 80 kDa, 34 kDa, 30 kDa, AND 29 kDa).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- DOMAIN: THE DNA POLYMERASE ACTIVITY DOMAIN RESIDES IN THE N-TERMINAL HALF OF THE PROTEIN, WHILE THE C-TERMINUS IS NECESSARY FOR COMPLEXING SUBUNITS B AND C.
 CC -!- MISCELLANEOUS: In eukaryotes there are five DNA polymerases: alpha, beta, gamma, delta, and epsilon which are responsible for different reactions of DNA synthesis.
 CC -!- SIMILARITY: Belongs to the DNA polymerase type-B family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M60416; AAA88711.1;
 CC EMBL: X92494; CAA63235.1;
 CC EMBL: Z71538; CAA96169.1;
 CC EMBL: A36028; A36028.
 CC GERMOnline: 143268.
 CC SGD: S0005206; POL2.
 CC GO: GO:0000731; P:DNA repair synthesis; IMP.
 CC InterPro: IPR006172; DNA_pol_B.
 CC InterPro: IPR006134; DNA_pol_B_dom.
 CC InterPro: IPR006133; DNA_pol_B_exo.
 CC Pfam: PF00136; DNA_pol_B; 1.
 CC SMART: SM00486; POLBc; 1.
 CC PROSITE: PS00116; DNA_POLYMERASE_B; FALSE NEG.
 CC TRANSFASER: DNA-directed DNA polymerase; DNA replication;
 CC DNA-binding; Zinc-finger; Nuclear protein.

FT ZN FING 2108 2181 POTENTIAL.
 FT VARIANT 644 644 M -> I (IN POL2-9 TS MUTANT).
 FT VARIANT 710 710 P -> S (IN POL2-18 TS MUTANT).
 SQ SEQUENCE 2222 AA; 255669 MW; CBCDDE2ABI47D65B CRC64;
 Query Match 6.8%; Score 85.5; DB 1; Length 2222;
 Best Local Similarity 28.2%; Pred.No. 56;
 Matches 37; Conservative 13; Mismatches 50; Indels 31; Gaps 7;
 QY 53 LPVPWP-TLVTTFSXGVQCFSRYPDHM-----KRHDFFKSAMPEGYV---QERTI 98
 DB 883 LPKSPFETFTLENGKKLLSLSPCSMLNVRVHQKFTNHQYQELKDPINTVIYETHSENTI 942
 QY 99 FFKDDGNYKTR--AEVFEFGDTLVNR-----IELKGIDPFKEDGNILGHKLEYNFI 146
 DB 943 FFEVDGPKAMILPSSKEGKIKRYAVNFNEDGSLAELKGFELKRGEL---QLINKFQ 999
 QY 147 S--HNVTAD 155
 DB 1000 SDIFKVFLEGD 1010
 RESULT 5
 AC2L HUMAN STANDARD; PRT; 689 AA.
 ID AC2L_HUMAN STANDARD; PRT; 689 AA.
 AC Q9NUI; Q8IV99; Q8N234; Q96J11; Q96UX6; Q9NU28;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Acetyl-coenzyme A synthetase 2-like, mitochondrial precursor
 DE (EC 6.2.1.1) (Acetate--CoA ligase 2) (Acetyl-CoA synthetase 2).
 GN ACAS2L OR KIAA1846.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R., Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L., Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M., Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J., Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M., Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R., Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M., Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Hawley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Levaeslaih M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.I., Martin S.I., McConachie L.J., McIlroy K., McMurray A.A., Milne S.A., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.J.C.T., Prachalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=Brain, and Testis;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

DR	ENBL; AK027817; BAB55390.1; ALT_INIT.
DR	ENBL; AK092295; BAC03853.1; ALT_SEQ.
DR	ENBL; AB058749; BAB47475.1; -
DR	Genew; HGNC:16091; ACAS2L
DR	InterPro; IPR000873; AMP-bind.
DR	Pfam; PF00501; AMP-binding; 1.
DR	PROSITE; PS00455; AMP BINDING; 1.
KW	Ligase; Mitochondrion; Transic peptide; Alternative splicing.
FT	TRANSIT 1 36
FT	CHAIN 37 689
FT	DOMAIN 45 53
FT	VARSPLIC 446 447
FT	CONFLICT 277 277
FT	CONFLICT 488 488
FT	SEQUENCE 689 AA; 74856 MW; 66E84E39302AD0B CRC64;
SQ	
 Query Match 6.7%; Score 85; DB 1; Length 689; Best Local Similarity 24.1%; Pred. No. 15; Matches 33; Conservative 15; Mismatches 37; Indels 52; Gaps 7	
QY	8 FTGVVPIILVELDGVNGHKFSVSGEGEDATYKGLTLKFCTTGKLPWPMTLVTFPSXG 67
Db	473 FFGIVFLMDKGSV-----VEGNSVGALCIS-----QAMPGMARTI--- 510
QY	68 VQCFSRYPDHMKRHDFKFSAMPEGYVQERTIFFKDGNKYTRA--EVKFEGLTVNRLE 124
Db	511 -----YGDHQRFVDAYFKAP-QY-----YFTGDGARTEGGYQITGRMDDVI--- 553
QY	125 LKGIDFKEDGNILGHKL 141
Db	554 -----NISGRL 560
 RESULT 6 AMPA WIGBR STANDARD; PRT; 501 AA. ID AMPA WIGBR Q8D295; AC Q8D295; AC DT 10-OCT-2003 (Rel. 42, Created) DT 10-OCT-2003 (Rel. 42, Last sequence update) DE 10-OCT-2003 (Rel. 42, Last annotation update) DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase) DE (LAP) [leucyl aminopeptidase]. OS FEPA OR WIGBR4590. GN Wiglesworthia glossinidia brevipalpis. OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; OC Enterobacteriaceae; Wiglesworthia. NCBI_TaxID=36870; RN [1]_TaxID=36870; RP SEQUENCE FROM N.A. RX MEDLINE=22297718; PubMed=12219091; RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M., RA Aksay S.; RL "Genome sequence of the endocellular obligate symbiont of tsetse flies, Wiglesworthia glossinidia." Nat. Genet. 32:402-407(2002). CC -!- FUNCTION: Presumably involved in the processing and regular turnover of intracellular proteins. Catalyzes the removal of unsubstituted N-terminal amino acids from various peptides (By similarity). CC CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa- -Xbb-, in which Xaa is preferably Leu, but may be other amino acids including Pro although not Arg or Lys, and Xbb may be Pro. CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity). CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity). CC -!- SIMILARITY: Belongs to peptidase family M17. CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/)	


```

STRH_STRPN
ID STRH_STRPN STANDARD; PRT; 1312 AA.
AC P49610;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Beta-N-acetylhexosaminidase precursor (EC 3.2.1.52).
GN STR OR SP0057.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12213;
RX MEDLINE=95238375; PubMed=7721787;
RA Clarke V.A., Platt N., Butters T.D.;
RT "Cloning and expression of the beta-N-acetylglucosaminidase gene from
RT Streptococcus pneumoniae. Generation of truncated enzymes with
RT modified aglycon specificity."
RL J. Biol. Chem. 270:8805-8814(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gunn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RA "Complete genome sequence of a virulent isolate of Streptococcus
RA pneumoniae."
RL Science 293:498-506(2001).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing N-
CC acetyl-D-hexosamine residues in N-acetyl-beta-D-hexosaminides.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -!- SIMILARITY: Belongs to family 20 of glycosyl hydrolases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L36923; AAC41450.1; --
CC EMBL; AE007323; AAK74246.1; --
CC PIR; A56390; A56390.
CC PIR; E95006; E95006.
CC TIGR; SP0057; --
CC InterPro; IPR001540; Glyco_hydro_20.
CC InterPro; IPR005877; Gpos_Ysirk_
CC InterPro; IPR001899; Gram_pos_anchor.
CC Pfam; PF00728; Glyco_hydro_20; 1.
CC Pfam; PF00745; Gram_pos_anchor; 1.
CC Pfam; PF04650; Ysirk_signal; 1.
CC TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
CC TIGRFAMs; TIGR01168; Ysirk_signal; 1.
CC PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
CC Hydrolase; Glycosidase; Cell wall; Peptidoglycan-anchor; Repeat;
KW Signal; Complete proteome.
FT SIGNAL 1 33
FT CHAIN 34 1284
FT PROPEP 1285 1312
FT DOMAIN 176 616
FT DOMAIN 621 1046
FT SITE 1281 1285
FT MOD_RES 1284 1284

```

```

FT CONFLICT 39 39 MISSING (IN REF. 1).
FT CONFLICT 69 69 V -> E (IN REF. 1).
FT CONFLICT 169 169 A -> E (IN REF. 1).
FT CONFLICT 617 617 Q -> L (IN REF. 1).
FT CONFLICT 1045 1045 V -> A (IN REF. 1).
FT CONFLICT 1161 1161 E -> K (IN REF. 1).
FT CONFLICT 1171 1171 C -> R (IN REF. 1).
FT CONFLICT 1267 1267 V -> A (IN REF. 1).
SQ SEQUENCE 1312 AA; 144549 MW; 503375B5257A90B5 CRC64;

Query Match 6.4%; Score 81; DB 1; Length 1312;
Best Local Similarity 20.7%; Pred.No. 71;
Matches 56; Conservative 36; Mismatches 101; Indels 78; Gaps 13;

Qy 21 DVNGHKFSVSGEGGDATYKGLTKFTCTGKLP-----VVPHTLVITFSXGV- 68
Db 896 NTNGDWYILGQKPEDG--GGFLKKAJENTGKTFNGLASTKYPEVDLPVGSMLSIWAD 953
Qy 69 -----QCFSRYPDHMKRHHFFK-----SAMP---EGYVOERTIFF--- 100
Db 954 RPSAEYKEEIFELMTAFADHNK--DYFRANYNALREELAKIPTNLEGYSKESLEALDAA 1011
Qy 101 KDCGNKYKTRAEVKFEGDTLVNRIE--LKGIDFKEDGNILGHKLEYNFISHNVITADKQK 158
Db 1012 KTAIYNLNLNENKQALDITLVANLKAALQGLK-----PAVTH-----SGSLDE 1053
Qy 159 NGIKANFKIRHNI-----EDGSVLADHYQONTPIGDGPVLLPD-----NHYLSTQSALSK 209
Db 1054 NEVAANVETPELITRTTEIPFEVKKENPNLPAQENIITAGVKGBERTHVISVLTENGK 1113
Qy 210 DPKEKRDHMLLEFV-----TAAGITHGMD 235
Db 1114 TTEVLDSQVTKVINOVEVGPVTHKGDE 1144

RESULT 14
CP51_CANGA
ID CP51_CANGA STANDARD; PRT; 533 AA.
AC PS0855; Q02312;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome P450 51 (EC 1.14.13.70) (CYPL1) (P450-LIAL) (Sterol 14-
DE alpha-demethylase) (Lanosterol 14-alpha demethylase) (P450-14DM).
DE ERG11 OR CYP51.
GN Candida glabrata (Yeast) (Torulopsis glabrata).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2001-L5;
RX MEDLINE=96161286; PubMed=8593007;
RA Geber A., Hitchcock C.A., Swartz J.E., Pullen F.S., Marsden K.E.,
RA Kwon-Chung K.J., Bennett J.E.;
RA "Deletion of the Candida glabrata ERG3 and ERG11 genes: effect on cell
RA viability, cell growth, sterol composition, and antifungal
RA susceptibility."
RL Antimicrob. Agents Chemother. 39:2708-2717(1995).
RN [2]
RP SEQUENCE OF 60-473 FROM N.A.
RC STRAIN=ATCC 2001;
RX MEDLINE=95081364; PubMed=7989540;
RA Burgence-Kauriz P., Zuber J.P., Jaunin P., Buchman T.G., Bille J.,
RA Rossier M.;
RA "Rapid detection and identification of Candida albicans and
RA Torulopsis (Candida) glabrata in clinical specimens by
RA species-specific nested PCR amplification of a cytochrome P-450
RA lanosterol-alpha-demethylase (Lial) gene fragment."
CC J. Clin. Microbiol. 32:1902-1907(1994).
CC -!- FUNCTION: Catalyzes C14-demethylation of lanosterol which is
CC critical for ergosterol biosynthesis. It transforms lanosterol
CC into 4,4'-dimethyl cholesta-8,14,24-triene-3-beta-ol (By

```



```

76 80
81 82
88 88
93 95
101 102
105 106
111 115
116 117
120 132
137 142
144 148
149 150
153 164
165 165
168 171
175 176
181 181
182 195
196 198
206 207
217 217
218 219
220 223
230 232
234 239
243 243
245 249
250 252
256 259
262 276
280 282
288 301
305 333
334 335
337 342
344 345
346 356
357 358
360 366
371 374
377 379
380 381
401 403
410 411
412 417
418 420
423 423
424 424
430 432
437 440
441 442
445 445
459 461
464 472
473 474
476 477
483 484
491 506
509 513
517 521
523 524
SQ SEQUENCE 533 AA; 58994 MW; D795B73D145F9154 CRC64;

```

Search completed: April 19, 2004, 15:06:28
Job time : 18 secs

```

283 CNFIGVDGIVETLRDMAKCFDD-PELTIR-TEEVIAEBEIAAQDDLDYFKEKL 333
111 EVK-----PEGDTLVNRIELKIDFKEDGNILGHKLEY-----NFISHN-- 149
334 QGKTACLYVGGSRSHTYNNMLKSFVGSLSVAGFEFAHRDDYEGREVIPITIKIDADSKNIP 393
150 -VYITADKQKNGIKANFKIRHNIEDGVSQVLADHYQQONTPIGDGPVLLPDNHNHLSQSAIS 208
394 EITVTPDEQKVRVVPEDKVBELKKAGVFLSSYGGMKEMHDGTTLIDDMNHHDMVEVVLE 453
209 K 209
454 K 454

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 15:03:20 ; Search time 20 Seconds
(without alignments)
1144.679 Million cell updates/sec

Title: US-10-057-505-2-COPY

Perfect score: 1265

Sequence: 1 MSKGEEFTGVVPILVLDG.....VLLEFVTAAGITHGMDELYK 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1228	97.1	238	1 J01514	Green-fluorescent
2	96	7.6	785	2 D72228	hypothetical prote
3	91.5	7.2	2573	2 D71614	hypothetical prote
4	88	7.0	281	2 AD2052	hypothetical prote
5	87.5	6.9	357	2 G81355	tRNA (uracil-5)-m
6	87.5	6.9	613	2 A99552	oligoendopeptidase
7	87	6.9	173	2 A83629	sterol binding pro
8	85.5	6.8	2222	1 A36028	DNA-directed DNA p
9	84.5	6.7	324	2 A83548	vegetable incomp
10	84.5	6.7	889	2 JC5576	inter-alpha-trypsi
11	84	6.6	529	2 B86815	hypothetical prote
12	83	6.6	632	2 T06586	DNA-binding protei
13	83	6.6	646	2 F71620	hypothetical prote
14	82.5	6.5	564	2 E81317	ABC-type transport
15	82.5	6.5	877	2 D7H4708	iron-regulated out
16	82.5	6.5	887	2 E82590	leucyl-tRNA synthe
17	82.5	6.5	1138	2 H86201	hypothetical prote
18	82	6.5	471	2 T27856	hypothetical prote
19	82	6.5	797	2 JC4078	protective surface
20	82	6.5	808	2 F84102	fibronectin-bindin
21	82	6.5	1092	2 D42798	polyketide synthas
22	81.5	6.4	2518	2 A12140	glucose dehydrogen
23	81	6.4	371	2 B83911	protein F1086.14 [
24	81	6.4	760	2 F86281	beta-N-acetylhexos
25	81	6.4	1312	2 E95006	hypothetical prote
26	80.5	6.4	322	2 T22410	iron-sulfur cofact
27	80.5	6.4	370	2 E70390	nitrogenase (EC 1.
28	80.5	6.4	534	1 N1CLWA	hypothetical prote
29	80.5	6.4	636	2 C70031	

ALIGNMENTS

RESULT 1

JQ1514

green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)

C:Species: Aequorea victoria

C:Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 23-Mar-2001

C:Accession: JQ0692; JQ1514; PQ0335; S48693; S51330; S51331

R:Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.

Gene 111, 229-233, 1992

A:Title: Primary structure of the Aequorea victoria green-fluorescent protein.

A:Reference number: JQ1514; MUID:92175527; PMID:1347277

A:Accession: JQ0692

A:Molecule type: DNA

A:Residues: 1-107,'S',109-238 <PRA1>

A:Cross-references: GB:M62654; NID:G155662; PIDN:AAA27722.1; PID:G155663

A:Accession: JQ1514

A:Molecule type: mRNA

A:Residues: 1-99,'F',101-140,'L',142-218,'V',220-238 <PRA2>

A:Cross-references: GB:M62653; NID:G155660; PIDN:AAA27721.1; PID:G155661

A:Accession: PQ0335

A:Molecule type: protein

A:Residues: 46-64,74-122,132-151,154-183,185-200 <PRA3>

R:Inouye, S.; Tsuji, F.I.

FEBS Lett. 351, 211-214, 1994

A:Title: Evidence for redox forms of the Aequorea green fluorescent protein.

A:Reference number: S48693; MUID:94364470; PMID:8082767

A:Accession: S48693

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-24,'Q',26-156,'P',158-171,'K',173-238 <INO>

A:Cross-references: GB:L29345; NID:G606383; PIDN:AAA58246.1; PID:G606384

R:Watkins, J.N.; Campbell, A.K.

submitted to the EMBL Data Library, January 1995

A:Reference number: S51330

A:Accession: S51330

A:Molecule type: mRNA

A:Residues: 1-13,'V',15-24,'Q',26-44,'N',46-153,'G',155-156,'P',158-171,'K',173-227,'R',

A:Cross-references: EMBL:X83959; NID:G634008; PIDN:CAA58789.1; PID:G634009

A:Experimental source: clone gfp1

A:Accession: S51331

A:Molecule type: mRNA

A:Residues: 1-24,'Q',26-29,'R',31-83,'L',85-153,'G',155-156,'P',158-171,'K',173-208,'Q',

A:Cross-references: EMBL:X83960; NID:G634010; PIDN:CAA58790.1; PID:G634011

A:Experimental source: clone gfp2

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

submitted to the Brookhaven Protein Data Bank, August 1996

A:Reference number: A6592; PDB:1GFL

A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A',2-79,'R',81-;

A>Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

Nat. Biotechnol. 14, 1246-1251, 1996

A:Title: The molecular structure of green fluorescent protein.

A:Reference number: A58953; MUID:98294543; PMID:9631087

A;Contents: annotation; X-ray crystallography, 1.9 angstroms
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFNV) emitting
F;65-67/Cross-link: 5-imidazolone (Ser-Gly) #status experimental
F;66/Modified site: dehydrotyrosine (Tyr) #status experimental
C;Genetics:

A;Gene: GPP
A;Introns: 69/3; 167/3
C;Superfamily: green-fluorescent protein
C;Keywords: chromoprotein; luminescence
F;65-67/Cross-link: 5-imidazolone (Ser-Gly) #status experimental
F;66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match 97.1%; Score 1228; DB 1; Length 238;
Best Local Similarity 95.8%; Pred. No. 5e-98; 5; Indels 0; Gaps 0;
Matches 228; Conservative 5; Mismatches 5;

QY 1 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGDATYKGLTLKFTCTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPLVELDGVNGHKFSVSGEGDATYKGLTLKFTCTTGKLPVWPPTL 60
QY 61 VTFSXGVQCFSRYPDHMKRHDFKSAPEGYVOERTIFFKDDGNYKTRAEVKEGDTLV 120
Db 61 VTFSXGVQCFSRYPDHMKRHDFKSAPEGYVOERTIFFKDDGNYKTRAEVKEGDTLV 120
QY 121 NRIELKIGIDFKEDGNILGHKLEYNFISHNVYITADKQNGI;KANFKIRHNIEDGVSQVAD 180
Db 121 NRIELKIGIDFKEDGNILGHKLEYNFISHNVYITADKQNGI;KANFKIRHNIEDGVSQVAD 180
QY 181 HYQNTPIGDGVLPPDNHLYSTQALSQKPKERDMMVLEFVTAAGITGMDELK 238
Db 181 HYQNTPIGDGVLPPDNHLYSTQALSQKPKERDMMVLEFVTAAGITGMDELK 238

RESULT 2

H72228
hypothetical protein TM1624 - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: H72228
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: H72228

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-785 <ARN>
A;Cross-references: GB:AE001806; GB:AE000512; NID:g4982196; PIDN:AA36691.1; PID:g498219
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1624

Query Match 7.6%; Score 96; DB 2; Length 785;
Best Local Similarity 20.9%; Pred. No. 3.5;
Matches 40; Conservative 25; Mismatches 66; Indels 60; Gaps 5;

QY 2 SKGEELFTGVVPLVELDGVNGHKFSVSGEGDATYKGLTLKFTCTTGKLPVWPPTLV 61
Db 15 NEGRFSFEGTVPGVVQAD-----LVKGLLPHFYVQM- 46
QY 62 TTFSGXGVQCFSRYPDHMKRHDFKSAPEGYVOERTIFFKDDGNYKTRAEVKEGDTLVN 121
Db 47 -----NEDLFKIEDREWIEREFKEKEDVKGERVDLVFEGVDTL 88
QY 122 RIELKIGIDFKEDGNILGHKLEYNFISHNVYITADKQNGI;KANFKIRHNIEDGVSQVAD 181
Db 89 DVLNGVYL---GSTEDMFIEVRFDVTNL-----KEKNHLKVYIK-----SPIRVPKT 134
QY 182 YQNTPIGDGP 192
Db 135 LEQNYGVLGPP 145

RESULT 3

D71614
hypothetical protein PFB0460c - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C;Accession: D71614

R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Peretea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.

A;Reference number: A71600; MUID:99021743; PMID:9804551
A;Accession: D71614

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2573 <GAR>

A;Cross-references: GB:AE001396; GB:AE001362; NID:g3845188; PIDN:AACT71881.1; PID:g38451;
A;Experimental source: clone 3D7
C;Genetics:
A;Gene: PFB0460c

Query Match 7.2%; Score 91.5; DB 2; Length 2573;
Best Local Similarity 26.9%; Pred. No. 41;
Matches 35; Conservative 31; Mismatches 51; Indels 13; Gaps 5;

QY 93 VQERTIFFKD--DQNYKTRAEVKEGDTLVNRIELKIGIDFKEDGNILGHKLEYNFISH-- 148
Db 126 LKXETILCKDIKSGNSDPMDIEISLFKDDMVDDKELK--DFEKSLSKTKNKKEVNFYINYM 183
QY 149 NVVITADKQNGI;KANFKIRHNIEDGVSQVADHYQNTPIGDGVLPPDNHLYSTQALS 208
Db 184 NLHLENKKKDEKKNKIHNNDNNM---IYYKNI---DKTHYILDNNVVHILNDIN 236

QY 209 KDPKPKRDHM 218

Db 237 TYLKRERYDM 246

RESULT 4

AD2052
hypothetical protein all1970 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AD2052
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AD2052
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-281 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA073669.1; PID:g17131060; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all1970

Query Match 7.0%; Score 88; DB 2; Length 281;
Best Local Similarity 25.2%; Pred. No. 4.5;
Matches 34; Conservative 27; Mismatches 52; Indels 22; Gaps 7;

QY 106 YKTRAEVKEGDTLVNRIELKGI----DFKEDGNILGHKLEYNFISHNVITADKQNGI 161
Db 46 FKARS-LQSSSDILLNPIKYNLNQRPDKQKQYIYHABGYK-KHE-----SSKRKHSI 99
QY 162 KANFKI--RHNIEDGVSQVADH-----YQNTPIGDG---PVLLPDNHLYSTQALS 209
Db 100 KPGFSLQERGDVNVQLLDNLNVLDPDPIQTDIPTFSGQNTPLAMRDYKFSQPKITK 159
QY 210 DPKPKRDHMVLEFV 224

R;Sen-Gupta, M.; Lyck, R.; Niedenthal, R.K.; Fleig, U.N.; Hegemann, J.H.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63235
A;Accession: S63235
A;Molecule type: DNA
A;Residues: 1-2222 <SEW>
A;Cross-references: EMBL:Z71538; NID:g1302316; PIDN:CAA96169.1; PID:g1302317; GSPDB:GN00146
A;Experimental source: strain S288C
R;Sen-Gupta, M.; Lyck, R.; Fleig, U.; Niedenthal, R.K.; Hegemann, J.H.
Yeast 12, 505-514, 1996
A;Title: The sequence of a 24 152 bp segment from the left arm of chromosome XIV from *Saccharomyces cerevisiae*
A;Reference number: S65111; MUID:96310631; PMID:8740425
A;Accession: S65121
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2221 <SEF>
A;Cross-references: EMBL:X92494; NID:g1045236; PIDN:CAA63235.1; PID:g1045247
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGB:POL2; DUN2; MIPS:YNL262W
A;Cross-references: SGD:S0005206; MIPS:YNL262W
A;Map position: 14L
C;Superfamily: DNA-directed DNA polymerase II
C;Keywords: DNA binding; nucleotidyltransferase; nucleus; zinc finger

Query Match 6.8%; Score 85.5; DB 1; Length 2222;
Best Local Similarity 28.2%; Pred. No. 1.1e+02;
Matches 37; Conservative 13; Mismatches 50; Indels 31; Gaps 7;

QY 53 LPVWP- TLVTFSGVQCFSKYPDHM-----KRHDFKAMPEGYV-----QRTI 98
DB 883 LPSKFPETYLENGKLLYSYFCMLNRYVHOKFTNHQYQELKDPINLYIYHSENTI 942
QY 99 FFKDDGNYKTR--AEVPEGDTLVNR-----IELKIDPKEDGNILGHKLEYNFI 146
DB 943 FFEVDGPKAMILPSSKEEGKIKRYAVFEDNGSLAELKGFELKRGEL---QLKNFQ 999
QY 147 S--HNVVITAD 155
DB 1000 SDIFKVFLEGD 1010

RESULT 9
AB3548
vegetable incompatibility protein het-e-1 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AB3548
R;DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3548
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-324 <KUP>
A;Cross-references: GB:AF008918; PIDN:AAL53549.1; PID:g17984458; GSPDB:GN00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10307
A;Map position: II

Query Match 6.7%; Score 84.5; DB 2; Length 324;
Best Local Similarity 25.7%; Pred. No. 11;
Matches 48; Conservative 19; Mismatches 61; Indels 59; Gaps 11;

QY 8 FTGVVPIVLVDGDV---NGHKFS-----VSGE---GEGDATYVGLTLKFI 47
DB 18 FIGGIPFFAMADCAVRLDDGGHQTSTVHDGLLSAAMVDGKSLVTGGEDGRVCRIDAKGV 77
QY 48 CT-TGKLFVPPVTLVTT-----FSXGVQCFSRYPEDHMKRHDFFKSAPEGYVOERTIF 99

DB 78 VTLANIPKWTAVATGPNGTGTFASGKTAWSRADG-KVQEFTOERSVEG-----IA 130
QY 100 FKDDGNYKTRAEVKFEGDTLV-----NRIEIK-----GIDFKEDGNIL----- 137
DB 131 FAPKG--QRLATARYNGATLIWAGTAAKPVLEWKAGHIGITFSPDGRFLITSMQENALH 188
QY 138 GHKLEYN 144
DB 189 GWRLEDN 195

RESULT 10
JC5576
inter-alpha-trypsin inhibitor heavy chain 3 - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 17-Mar-1999
A;Accession: JC5576; PC4486
R;Nakatani, T.; Suzuki, Y.; Yamamoto, T.; Sinohara, H.
J. Biochem. 122, 71-82, 1997
A;Title: Molecular cloning and sequencing of cDNAs encoding three heavy-chain precursors of inter-alpha-trypsin inhibitor heavy chain family.
A;Reference number: JC5574; MUID:97420688; PMID:9276673
A;Accession: JC5576
A;Molecule type: mRNA
A;Residues: 1-889 <NAK>
A;Cross-references: DDBJ:D89287
A;Experimental source: liver
A;Accession: PC4486
A;Molecule type: protein
A;Residues: 34-53; 449-475; 509-526 <NA2>
C;Comment: In the plasma three inter-alpha-trypsin inhibitor heavy chains 1, 2 and 3 exist that the complexes play important role for pancreatic cancer.
C;Superfamily: inter-alpha-trypsin inhibitor complex component II
F;236-239,664-865/Disulfide bonds: #status predicted

Query Match 6.7%; Score 84.5; DB 2; Length 889;
Best Local Similarity 24.5%; Pred. No. 41;
Matches 37; Conservative 33; Mismatches 52; Indels 19; Gaps 7;

QY 74 YPDHMKRHDFFKSAMPEGYVOERT-----IFFKDDGNYKTRAEVKFEGD-----TLVNRIE 124
DB 479 YPENAIL-DLTNKSYPHFYDGETAVAGRLADSDMMNFK--ADVKGHGALNDLTFTTEED 535
QY 125 LKGD--FKEDGNILGHKLEYNVISHNVITADKNGIKANFKIRNIEDSGSVQLADHY 182
DB 536 MKENDAAALKEQGYIFGNYIERLWAYLTITOLLEKRN---AHGEEKENLTAQALELSIKY 592
QY 183 QQNTPIGDGPFVLLPDNDHYLSTQSALSXPKE 213
DB 593 HFVTFPLPMVTKPEDN--EDQTSIADKPGE 621

RESULT 11
B86815
hypothetical protein ypgd [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: B86815
R;Bolotin, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich, Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s.
A;Reference number: B86815; MUID:21235186; PMID:11337471
A;Accession: B86815
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-529 <STO>
A;Cross-references: GB:AE005176; PID:g12724521; PIDN:AAK05620.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ypgd

Query Match 6.6%; Score 84; DB 2; Length 529;
Best Local Similarity 22.3%; Pred. No. 23;

C;Species: Helicobacter pylori
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 29-Sep-1999
 C;Accession: H64708
 R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A;Reference number: A64520; MUID:97394467; PMID:9252185
 A;Accession: H64708
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-877 <TOM>
 A;Cross-references: GB:AE000649; GB:AE000511; NID:g2314687; PIDN:AAD08553.1; PID:g231469
 C;Superfamily: iron-regulated outer membrane protein

Query Match	6.5%;	Score 82.5;	DB 2;	Length 877;
Best Local Similarity	35.3%;	Pred. No. 59;		
Matches	36;	Conservative	7;	Mismatches 26; Indels 33; Gaps 8;

Qy	115	EGDTLVNRIEL--KGIDFK-----EDGNILGHKLEYNFISHN--VYITADKQKNGIKAN	164
Db	454	EG-TLARRIFLNSGVNFKVTHPISED---YGNVFEXGMIVQNLSVPSGLDKGKNGYYKN	509
Qy	165	FKIRHNIEDGSVQLADHYQQNTPIGDPVLLPDNHYLSTQSA	206
Db	510	-----NID-----PNDPNGPG---LPYRHYTDQSS	532

Search completed: April 19, 2004, 15:07:59
 Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2004, 15:07:31 ; Search time 47 Seconds
(without alignments)
1395.962 Million cell updates/sec

Title: US-10-057-505-2-COPY
Perfect score: 1265
Sequence: 1 MSKGEELFTGVVPIVLDG.....VLLEFVTRAGITHGMDLYK 238

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 1124875 seqs, 275673149 residues

Total number of hits satisfying chosen parameters: 1124875

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	98.0	545	14	US-10-214-932-52
2	1240	98.0	548	14	US-10-214-932-76
3	1240	98.0	1192	14	US-10-214-932-54
4	1238	97.9	238	9	US-09-884-681-2
5	1238	97.9	238	10	US-09-967-301-2
6	1238	97.9	238	12	US-10-457-982-2
7	1238	97.9	238	13	US-10-024-686-2
8	1238	97.9	238	13	US-10-057-505-2
9	1238	97.9	238	14	US-10-293-580-2
10	1238	97.9	238	14	US-10-293-580-74
11	1235	97.6	387	11	US-09-327-876-87
12	1235	97.6	387	15	US-10-360-149-87
13	1234	97.5	238	9	US-09-920-922-4
14	1234	97.5	238	9	US-09-852-000-1
15	1234	97.5	238	10	US-09-900-345A-125

16	1234	97.5	238	10	US-09-866-538-2	Sequence 2, Appli
17	1234	97.5	238	10	US-09-794-308-2	Sequence 2, Appli
18	1234	97.5	238	10	US-09-865-291-2	Sequence 2, Appli
19	1234	97.5	238	12	US-10-132-067-8	Sequence 8, Appli
20	1234	97.5	238	14	US-10-121-258-10	Sequence 10, Appli
21	1234	97.5	238	14	US-10-221-461-6	Sequence 6, Appli
22	1234	97.5	238	14	US-10-305-765-10	Sequence 10, Appli
23	1234	97.5	238	14	US-10-305-765-159	Sequence 159, App
24	1234	97.5	238	14	US-10-305-633-10	Sequence 10, Appli
25	1234	97.5	238	14	US-10-305-633-159	Sequence 159, App
26	1234	97.5	238	15	US-10-370-570-1	Sequence 1, Appli
27	1234	97.5	238	15	US-10-370-570-3	Sequence 3, Appli
28	1234	97.5	238	15	US-10-370-570-53	Sequence 53, Appli
29	1233	97.5	238	9	US-09-899-954B-2	Sequence 2, Appli
30	1233	97.5	238	15	US-10-370-570-2	Sequence 2, Appli
31	1231	97.3	238	12	US-10-132-067-2	Sequence 2, Appli
32	1231	97.3	239	15	US-10-442-148A-3	Sequence 3, Appli
33	1231	97.3	243	10	US-09-900-345A-60	Sequence 60, Appli
34	1231	97.3	243	10	US-09-900-345A-62	Sequence 62, Appli
35	1231	97.3	243	10	US-09-900-345A-64	Sequence 64, Appli
36	1231	97.3	243	10	US-09-900-345A-66	Sequence 66, Appli
37	1231	97.3	243	10	US-09-900-345A-68	Sequence 68, Appli
38	1231	97.3	243	10	US-09-900-345A-70	Sequence 70, Appli
39	1231	97.3	243	14	US-10-305-765-94	Sequence 94, Appli
40	1231	97.3	243	14	US-10-305-765-96	Sequence 96, Appli
41	1231	97.3	243	14	US-10-305-765-98	Sequence 98, Appli
42	1231	97.3	243	14	US-10-305-765-100	Sequence 100, App
43	1231	97.3	243	14	US-10-305-765-102	Sequence 102, App
44	1231	97.3	243	14	US-10-305-765-104	Sequence 104, App
45	1231	97.3	243	14	US-10-305-765-110	Sequence 110, App

ALIGNMENTS

RESULT 1
US-10-214-932-52
; Sequence 52, Application US/10214932
; Publication No. US20030100707A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: APB02/US
; CURRENT APPLICATION NUMBER: US/10/214,932
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Sequence
US-10-214-932-52

Query Match	98.0%	Score	1240	DB	14	Length	545
Best Local Similarity	97.9%	Pred. No.	4e-124				
Matches	233	Conservative	1	Mismatches	4	Indels	0
						Gaps	0
Qy	1	MSKGEELFTGVVPIVLDG	VN	GH	KF	SV	SG
Db	308	MSKGEELFTGVVPIVLDG	VN	GH	KF	SV	SG
Qy	61	VTTFSGVQCFSRYPDHMKHDF	FK	SA	MP	EG	YV
Db	368	VTTFSGVQCFSRYPDHMKHDF	FK	SA	MP	EG	YV
Qy	121	NRIELKIDPKDGNILGHKLE	YN	FI	SH	NY	II
Db	428	NRIELKIDPKDGNILGHKLE	YN	FI	SH	NY	II

Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFI SHNVYITADKQKNGIKANFKIRHNIEDGVSQVLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFI SHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 5

US-09-967-301-2
; Sequence 2, Application US/09967301
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Stubbs, Simon L.
; APPLICANT: Jones, Anne E.
; APPLICANT: Michael, Nigel P.
; APPLICANT: Thomas, Nicholas
; TITLE OF INVENTION: Fluorescent Proteins
; FILE REFERENCE: PA0111
; CURRENT APPLICATION NUMBER: US/09/967,301
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: GB 0109858.1
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-967-301-2

Query Match 97.9%; Score 1238; DB 10; Length 238;
Best Local Similarity 97.5%; Pred. No. 2e-124;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTCLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTCLKFICTTGKLPVPWPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFI SHNVYITADKQKNGIKANFKIRHNIEDGVSQVLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFI SHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 6

US-10-457-982-2
; Sequence 2, Application US/10457982
; Publication No. US20030212265A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Llopis, Juan
; APPLICANT: Wächter, Rebekka M.
; APPLICANT: Remington, S. James
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; FILE REFERENCE: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/10/457,982
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US/09/602,641
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/172,063

; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-457-982-2

Query Match 97.9%; Score 1238; DB 12; Length 238;
Best Local Similarity 97.5%; Pred. No. 2e-124;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTCLKFICTTGKLPVPWPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTCLKFICTTGKLPVPWPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKEGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFI SHNVYITADKQKNGIKANFKIRHNIEDGVSQVLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFI SHNVYIMADKQKNGIKVNFKIRHNIEDGVSQVLAD 180
QY 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 7

US-10-024-686-2
; Sequence 2, Application US/10024686
; Publication No. US20020123113A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; TITLE OF INVENTION: MODIFIED GREEN FLUORESCENT PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/024,686
; FILING DATE: 17-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/057,995
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/727,452
; FILING DATE: 10-OCT-1996
; APPLICATION NUMBER: US95/14692
; FILING DATE: 13-NOV-1995
; APPLICATION NUMBER: 08/337,915
; FILING DATE: 10-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/032002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-024-686-2

Query Match
Best Local Similarity 97.9%; Score 1238; DB 13; Length 238;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLYNFIHNHVIITADKQNGIKANPKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLYNFIHNHVIITADKQNGIKANPKIRHNIEDGSVQLAD 180
QY 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 8
US-10-057-505-2
; Sequence 2, Application US/10057505
; Publication No. US20020164674A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: AURORA BIOSCIENCES CORPORATION
; APPLICANT: TSJEN, Roger
; APPLICANT: HEIM, Roger
; APPLICANT: CUBITT, Andrew
; TITLE OF INVENTION: TANDEM FLUORESCENT PROTEIN CONSTRUCTS
; FILE REFERENCE: REGEN260-3
; CURRENT APPLICATION NUMBER: US/10/057,505
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 08/792,553
; PRIOR FILING DATE: 1997-01-31
; PRIOR APPLICATION NUMBER: US 09/396,003
; PRIOR FILING DATE: 1999-09-13
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-10-057-505-2

Query Match
Best Local Similarity 97.9%; Score 1238; DB 13; Length 238;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLYNFIHNHVIITADKQNGIKANPKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLYNFIHNHVIITADKQNGIKANPKIRHNIEDGSVQLAD 180
QY 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
```

```
RESULT 9
US-10-293-580-2
; Sequence 2, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/10/293,580
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US/09/129,192
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea
US-10-293-580-2

Query Match
Best Local Similarity 97.9%; Score 1238; DB 14; Length 238;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
QY 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFGDTLV 120
Db 61 VTTFSGVQCFSRYPDHMKRHDFFKSAMPEGVVOERTIFFKDDGNYKTRAEVKFGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLYNFIHNHVIITADKQNGIKANPKIRHNIEDGSVQLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLYNFIHNHVIITADKQNGIKANPKIRHNIEDGSVQLAD 180
QY 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQNTPTIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 10
US-10-293-580-74
; Sequence 74, Application US/10293580
; Publication No. US20030170767A1
; GENERAL INFORMATION:
; APPLICANT: Aurora Biosciences Corporation
; APPLICANT: Cubitt, Andrew B.
; TITLE OF INVENTION: Fluorescent Protein Sensors of Post-Translational Modifications
; FILE REFERENCE: AURO1270 (08366/031001)
; CURRENT APPLICATION NUMBER: US/10/293,580
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US/09/129,192
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 74
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea green fluorescent protein phosphorylation mutant
US-10-293-580-74

Query Match
Best Local Similarity 97.9%; Score 1238; DB 14; Length 238;
Matches 232; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGLTLKFICTTGKLPVWPPTL 60
```

```
QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 180
QY 181 HYQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 238
Db 181 HYQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 238

RESULT 11
US-09-927-876-87
; Sequence 87, Application US/09927876
; Publication No. US2004000554A1
; GENERAL INFORMATION:
; APPLICANT: El Tayar, Nabil
; APPLICANT: Campbell, Robert K
; APPLICANT: Kelton, Christie A
; APPLICANT: He, Chaomei
; TITLE OF INVENTION: No. US2004000554A1el Glycoproteins and Methods of Use Thereof
; FILE REFERENCE: 20993-003
; CURRENT APPLICATION NUMBER: US/09/927,876
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/225,035
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/202,724
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
US-09-927-876-87

Query Match 97.6%; Score 1235; DB 11; Length 387;
Best Local Similarity 97.9%; Pred. No. 8.4e-124;
Matches 232; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 61
Db 151 SKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 210
QY 62 TTFPSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 121
Db 211 TTFPSYGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 270
QY 122 RIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 181
Db 271 RIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 330
QY 182 YQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 238
Db 331 YQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 387

RESULT 12
US-09-927-876-87
; Sequence 87, Application US/10360149
; Publication No. US20030219786A1
; GENERAL INFORMATION:
; APPLICANT: El Tayar, Nabil
; APPLICANT: Campbell, Robert K
; APPLICANT: Kelton, Christie A
; APPLICANT: He, Chaomei
; TITLE OF INVENTION: No. US20030219786A1el Glycoproteins and Methods of Use Thereof
; FILE REFERENCE: 20993-003
; CURRENT APPLICATION NUMBER: US/10/360,149

Query Match 97.6%; Score 1235; DB 11; Length 387;
Best Local Similarity 97.9%; Pred. No. 8.4e-124;
Matches 232; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 61
Db 151 SKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 210
QY 62 TTFPSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 121
Db 211 TTFPSYGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 270
QY 122 RIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 181
Db 271 RIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 330
QY 182 YQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 238
Db 331 YQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 387
```

```
; CURRENT FILING DATE: 2003-02-06
; PRIOR APPLICATION NUMBER: US/09/927,876
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/225,035
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/202,724
; PRIOR FILING DATE: 2000-05-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion Protein
US-10-360-149-87

Query Match 97.6%; Score 1235; DB 15; Length 387;
Best Local Similarity 97.9%; Pred. No. 8.4e-124;
Matches 232; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 SKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 61
Db 151 SKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 210
QY 62 TTFPSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 121
Db 211 TTFPSYGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 270
QY 122 RIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 181
Db 271 RIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 330
QY 182 YQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 238
Db 331 YQNTPIGDGPVLLPDNHYLSTQSALSQKPKRKHDMVLLFEVTAAGITHGMDELYK 387

RESULT 13
US-09-920-922-4
; Sequence 4, Application US/09920922
; Patent No. US20020083488A1
; GENERAL INFORMATION:
; APPLICANT: Miyawaki, Atsushi
; APPLICANT: Sawano, Asako
; TITLE OF INVENTION: METHOD FOR MUTAGENESIS
; FILE REFERENCE: 11283-012001
; CURRENT APPLICATION NUMBER: US/09/920,922
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: JP 2000-237166
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
US-09-920-922-4

Query Match 97.5%; Score 1234; DB 9; Length 238;
Best Local Similarity 97.1%; Pred. No. 5.3e-124;
Matches 231; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 60
Db 1 MSKGEELFTGVVPIILVELDGVNGHKFSVSGEGDATYGKLTLEFCTTGKLPVWPPTLV 60
QY 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSYGVQCFSRYPDHMKRHDFFKSAPEGYVOERTIFFKDDGNYKTRAEVKFEGDTLV 120
QY 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNFISHNVYITADKQKNGIKANFKIRHNIEDGVSQVAD 180
```

Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQNGIKVNFKIRHNIEDGSVOLAD 180
Qy 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 14

US-09-852-000-1
; Sequence 1, Application US/09852000
; Patent No. US20020099170A1
; GENERAL INFORMATION:
; APPLICANT: Osumi, Takashi
; APPLICANT: Tsukamoto, Toshiro
; APPLICANT: Tsukamoto, No. US20020099170A1iyo
; APPLICANT: Yanasaki, Masatoshi
; TITLE OF INVENTION: GREEN FLUORESCENT PROTEINS AND BLUE FLUORESCENT
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: 046124-5005-US
; CURRENT APPLICATION NUMBER: US/09/852,000
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: JP 026418/1998
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: US 09/121,539
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: US 09/615,655
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria
; FEATURE:
; OTHER INFORMATION: Green fluorescent protein
US-09-852-000-1

Query Match 97.5%; Score 1234; DB 9; Length 238;
Best Local Similarity 97.1%; Pred. No. 5.3e-124;
Matches 231; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MSKGEEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYKLTLPKFICTTGKLPVPWPTL 60
Db 1 MSKGEEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYKLTLPKFICTTGKLPVPWPTL 60
Qy 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQNGIKANFKIRHNIEDGSVOLAD 180
Qy 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

RESULT 15

US-09-900-345A-125
; Sequence 125, Application US/09900345A
; Publication No. US20030031999A1
; GENERAL INFORMATION:
; APPLICANT: Frazer, Ian Hector
; APPLICANT: Zhou, Jian
; TITLE OF INVENTION: METHOD AND POLYNUCLEOTIDES FOR DETERMINING TRANSLATIONAL
; TITLE OF INVENTION: EFFICIENCY OF A CODON
; FILE REFERENCE: 10338-5US
; CURRENT APPLICATION NUMBER: US/09/900,345A
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: AU PP8078
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: PCT/AU00/00008

; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 125
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: GFP humanized
; OTHER INFORMATION: control
US-09-900-345A-125

Query Match 97.5%; Score 1234; DB 10; Length 238;
Best Local Similarity 97.1%; Pred. No. 5.3e-124;
Matches 231; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 MSKGEEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYKLTLPKFICTTGKLPVPWPTL 60
Db 1 MSKGEEELFTGVVPIILVELDGVNGHKFSVSGEGEDATYKLTLPKFICTTGKLPVPWPTL 60
Qy 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Db 61 VTTFSXGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLV 120
Qy 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQNGIKANFKIRHNIEDGSVOLAD 180
Db 121 NRIELKGIDFKEDGNILGHKLEYNNSHVYIMADKQNGIKANFKIRHNIEDGSVOLAD 180
Qy 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238
Db 181 HYQONTPIGDGPVLLPDNHYLSTQSALS KDPKPKRDHMLLEFVTAAGITHGMDELYK 238

Search completed: April 19, 2004, 15:13:19
Job time : 47 secs